

The Effect of Test Delivery Mode (Online Vs. Paper-Based) On EFL Integrated Writing Assessment

Marziye Keshavarz (marzi.kz1400@gmail.com)

Department of English language and literature, Faculty of Humanities, Persian Gulf University, Bushehr 75169, Iran

Nasim Ghanbari (btghanbari@gmail.com) *Corresponding Author

Department of English language and literature, Faculty of Humanities, Persian Gulf University, Bushehr 75169, Iran

Parisa Abdolrezapour (abdolrezapour@gmail.com)

Department of Foreign Languages and Linguistics, Shiraz University, Shiraz, Iran

Abstract

In response to the increasing demand for online delivery of language tests, different international and local language testing bodies have considered the feasibility of delivering paper-based tests online. Integrating writing tasks has received increasing attention in diverse EFL/ESL settings. However, compared with the writing-only tasks, they are new practices in the Iranian EFL context. Despite many studies comparing the integrated and writing-only writing tasks, no studies have investigated the effect of the mode of delivery (online vs. paper-based) on the integrated writing performance of the test-takers in the Iranian EFL context. Adopting a quasi-experimental research approach, 44 Iranian EFL male and female students were divided into two online and paper-based groups. The participants received integrated writing instruction for the whole semester (16 weeks). At the end of the course, the students' integrated writing samples were rated holistically using the TOEFL integrated rubric. The results of analysis using independent samples T-tests showed that there was a significant difference between the two groups. The study discusses that when composing a text, online modality reduces the learners' psychological burden, improving the quality of integrated texts produced by the students. The pedagogical implications for EFL writing assessment are discussed.

Keywords: Integrated writing, Integrated writing assessment, EFL writing assessment, online assessment, paper-based assessment, mode of test delivery

Introduction

Integrated writing tasks have gained popularity recently and have been widely used to measure second-language learners' writing competency (Knochka & Sitajalabhorn, 2013). Integrated writing tasks encompass various definitions according to different researchers (e.g., Delaney, 2008; Cumming, 2011; Esmaeili, 2002; Gebril, 2010; Plakans, 2008, 2009a, 2009b, 2010; Plakans & Gebril, 2012; Yang, 2009; Yu, 2013); however, the explanation provided by Cumming et al. (2005) is the most explicit definition of integrated writing tasks. They claim that test takers must create written content that incorporates evidence from sources meaningfully in integrated tasks. This encompasses

understanding, integrating, and presenting ideas from the sources while also adhering to the appropriate style and protocols for citation and crediting the sources.

According to Cumming et al. (2007), the Educational Testing Service (ETS) introduced the integrated task on the TOEFL exam. This action aimed to assess test takers' writing skills more accurately. In addition, it was intended to assist individuals in enhancing their writing abilities and ensure that individuals can write appropriately in academic contexts (Barkaoui, 2015). Similarly, other authors claimed that integrated writing tasks focus on the truthfulness and validity and fairness of the test and offer test takers language and content for their writing (Cumming, et al., 2004; Feak & Dobson, 1996; Leki & Carson, 1997; Plakans & Gebril, 2012; Raimes, 1998; Weigle, 2004). In academic writing, numerous experts and researchers have abandoned traditional testing methods and now advocate for incorporating sources. In addition, most research on integrated essay writing tasks has centered on assessing second language writing abilities rather than emphasizing writing instruction (Gholami & Alinasab, 2017).

Ghanbari and Abdolrezapour (2021) argue that students struggle to write a composition based solely on reading. It is a complicated process involving the simultaneous incorporation of many elements (Flower & Hayes, 1981). Besides language abilities such as reading, spelling, grammar, and structuring paragraphs (Risemberg, 1996), students composing integrated writing assignments confront additional non-linguistic obstacles. The writer is participating in tasks that necessitate significant motivation and cognitive efforts. Writers' approaches to their writing tasks can be influenced by factors such as their goals, beliefs, attitudes, and predispositions (Ghanbari & Abdolrezapour, 2021).

Assessment plays a vital role in all teaching programs. It observes the standards of education and the extent to which students acquire knowledge. Furthermore, it assesses the students' level of achievement and the extent to which they have accomplished the curriculum objectives. Assessment activities ensure that what students have learned corresponds with the educational program's desired learning objectives (Ghanbari & Nowroozi, 2021). As defined by Webber (2012, p. 202), assessment refers to activities that aid student learning. Teachers need to have the capacity to develop and utilize assessments as a means to foster student growth. Incorporating technology in assessing students' learning has improved learning outcomes (Elmahdi et al., 2018) and a more efficient learning experience (Mohammadi, 2018).

Literature review

Theoretical backgrounds

Independent writing tasks have traditionally been used in academic contexts for instruction and assessment purposes. However, independent tests have been criticized on several grounds, more notably for their authenticity and construct validity (Cumming, 1997; Lumley, 2005; Weigle, 2002). As a result, the language testing community proposed integrated writing test tasks within which all the tasks are thematically interrelated. The source input provided the origin of the response(s) to be produced by the examinees (Lewkowicz, 1997, p. 121). The construction-integration model of reading comprehension (Kintsch, 1998) provides the theoretical background for the integrated writing tasks. Based on this model, understanding and production are mutually related processes, and both are dynamically involved in integrated writing tasks. The model states that production consists of constructing multiple cognitive representations at different propositional, textual, and situational levels (Cumming, 2013).

Compared to independent test tasks, integrated writing tasks are more authentic and valid (Luoma, 2004; Plakans, 2008; Weigle, 2004) and develop equity and fairness in academic settings (Huang & Hung, 2010). Providing reading input in a writing test can better simulate a standard academic writing task in which students are expected to write after reading several passages (Braine, 1989; Carson, 2001; Horowitz, 1986). Although there is no standard and conclusive model of L2 reading-writing associations (Cumming, 1997; Hirvela, 2016), pieces of evidence have been offered by various studies showing that reading-to-write tasks promote validity and are of the learners' interest and preference over independent tasks (Plakans, 2008; Soltani & Kheirzadeh, 2017). Considering fairness integrated writing tasks provide rich input for the lower ability learners and reduce the negative effects caused by the complexity of the task (Ghanbari & Abdolrezaapour, 2021; Weigle, 2004). Ultimately, integrated writing confidently guides language teachers toward a more holistic approach to teaching (Miller & Legg, 1993). As a considerable amount of time is spent teaching high-stakes tests such as IELTS and TOEFL, introducing integrated tasks paves the ground to pay equal attention to all language skills involved. This makes the language assessment more real-life and consequently more valid (Ghanbari & Abdolrezaapour, 2021).

In recent years, there has been an increased emphasis on integrated writing assessment (Gholami & Alinasab, 2017). Indeed, integrated writing tasks have been implemented by examiners to evaluate academic writing abilities in standardized proficiency tests. As mentioned earlier, the examinees are expected to extract information from the listening and reading input provided and synthesize it in a written mode (Gholami & Alinasab, 2017). Both large-scale tests, such as the Test of English as a Foreign Language (TOEFL), and classroom assessments use integrated assessment to evaluate test-takers writing abilities (Plakans et al., 2019).

Online assessment has been a new experience for many teachers and students. Several investigations have shown that online evaluation improves student learning outcomes (Dwiyanti & Suwastini, 2021). Gikandi et al. (2011) claim that students generally prefer online tests. In a study by Zlatovic et al. (2015), the online assessment of the students resulted in a change in their learning strategies to align with the expected ones, helping them achieve their learning goals.

Experimental Backgrounds

Numerous studies conducted in recent years highlight the increasing adoption of digital technology in English pedagogy across various EFL/ESL settings (Abdolrezaapour & Ghanbari, 2022; Ghanbari & Nowroozi, 2021; Hafner & Miller, 2011; Mompean, 2010; Ros et al., 2010); Rahimi et al., (2020) believe that computers provide novel and unexplored solutions to fulfill the requirements arising from EFL writing strategies and frameworks. Established computer software utilized for language learning can be viewed as reliable instruments for educators and students alike, facilitating the improvement of their writing skills, offering support to those learning English as a foreign language, and aiding them in overcoming the limitations of traditional methods. This helps them get better at writing.

Several studies have investigated the perspectives of Iranian students regarding utilizing computers for language learning. (e.g., Abdolrezaapour, 2019; Marandi, 2002). However, only a limited number of studies have examined the use of online writing assessment programs (Rahimi et al., 2020). For example, Marandi and Nami (2012)

investigated the impact of online writing tools on improving the coherence of English essays of Iranian female learners. The results showed that students who opted for computers as their writing tool exhibited a higher frequency of coherence indicators than their counterparts who relied on books.

Existing literature documents some studies which have investigated the mode of test delivery (paper-based vs. online) on the outcome and composing processes of the texts (Barkaoui & Knouzi, 2018; Barkaoui, 2016; Blackhurst, 2005; Breland et al., 2004; Green & Maycock, 2004; Guapacha Chamorro (2022); Jin & Yan, 2017; Lee, 2002; Li, 2006; Liu & Liu, 2023; Russell & Haney, 1997; Weir, et al., 2007; Wolfe & Manalo, 2005). These studies aimed to determine whether scores on computer-based tests (CBT) demonstrate equivalent ability to scores on similar paper-based tests (PBT) (Chapelle & Douglas, 2006). For example, many studies have compared PBT and CBT versions of the IELTS and TOEFL writing tests. Blackhurst (2005), Wolfe & Manalo (2005), and Weir (2007) concluded that there were no significant differences in writing styles across various approaches. In contrast, Green and Maycock (2004) discovered that test-takers exhibited a slight performance improvement when using the paper-based rather than the computer-based format. Breland et al.'s (2004) study found that although the average scores did not significantly vary across different writing modalities, a closer analysis of individuals with comparable English language abilities consistently showcased the superiority of the paper-based test. However, according to a recent study by Jin and Yan (2017), Chinese students achieved significantly higher scores when composing in English using computers instead of paper.

There is limited research on how the mode of test delivery (online vs. paper-based) can affect the quality of the integrated texts produced. For instance, Wolfe et al. (1996) and Whithaus et al. (2008) found that students' computer-based essays tended to be more organized, lengthier, and exhibited a more formal tone; however, the mode of test did not exert any influence on the frequency of mechanical errors of the texts. Russell and Haney (1997) found that the amount of writing produced by students who used computers for their essays was nearly doubled compared to those who used paper. Additionally, they tended to arrange their thoughts in multiple paragraphs. The language test Chambers (2008) administered involved evaluating two distinct groups. One group opted for computer-based essay writing; the other chose the conventional paper method. In contrast to paper responses, she found that computer responses contained a greater variety of words yet had fewer sentences and paragraphs. Nevertheless, no significant variations were observed in the essays in terms of length, vocabulary usage, punctuation, capitalization, and errors in spelling.

The findings above suggest that the writing modes can affect the writing processes used by test-takers, the characteristics of the texts they produce, and the scores they receive. If studies like Jin and Yan (2017) indicate that using a computer for writing leads to longer, more grammatically and vocabulary advanced, and more precise texts, then it is probable that essays composed on a computer will obtain higher scores than handwritten essays.

Similarly, the level of familiarity with writing on the computer appears to impact how test performance is influenced by writing on the computer. A study by Jin and Yan (2017) assessed individuals' performance on writing tasks. The results indicated that individuals' task performance was influenced by their familiarity with computers and the writing mode they used, either paper or computer. According to a study by Barkaoui & Knouzi

(2018), individuals with a greater understanding of computers generally perform better in computer assessments. Also, when people know more about using computers, their writing scores tend to be higher.

Yu and Iwashita (2021) compared the performance of a group of Chinese undergraduates on PBT and CBT. The findings revealed no significant difference in test scores between the modes. In addition, the computer familiarity survey showed that participants were generally familiar with using a computer. The attitude survey also demonstrated that half of the participants were positive towards CBT. Choi (2022) examined whether computer familiarity can be a potential source of construct-irrelevant variance in computer-based writing (CBW) tests. The study found that most test-takers held self-confidence and preferred CBW tests for reasons related to previous CBT writing experience (e.g., familiarity with CBT writing, useful tools/functions available on computers) regardless of L2 writing ability.

Guapacha Chamorro (2022) conducted a mixed-methods study to seek the cognitive validity evidence of CBT and PBT writing tests and their impact on EFL university students' cognitive processes (e.g., macro-planning, micro-planning), which determine the validity and fairness of classroom writing assessment. The findings showed that CBTs and PBTs were cognitively valid since both test modes could activate cognitive processes necessary to perform a writing assessment task. However, each affected EFL students' cognitive processes differently. While the PBT mode induced greater detailed macro-planning of content and text organization, the CBT mode triggered micro-planning of text organization and post-writing revisions. Also, the study showed no significant differences in task representation and translating processes for the two tests.

In another study, MovahedFar et al. (2022) investigated the effect of computer-based dynamic assessment (CBDA) on Iranian EFL learners' writing and attitude toward CBDA. The study found that teaching and assessing writing using computers can enhance the students' performance in writing.

Holland (1996) believes that the most predominant advantage of using computer software in writing is that editing and revising is much easier and simpler than revising on pencil and paper texts. Generally, all the students, including kids or adults, are enthusiastic about using word processors for writing, and their attitudes toward writing improve when required to write with word processors (Cochran-Smith, 1991). According to Wolfe et al. (1996), this positive attitude toward using computers for writing originates from word processors' editing and revising amenities. Similarly, Daiute (1986) and Cochran-Smith (1991) believe that editing tools offered to students make them convenient and encourage them to use word processors.

The Current Study

Over the past 50 years, technology and language education have been closely intertwined, and the utilization of computers in language learning has been a subject of study for around 30 years (Rahimi et al., 2020). During the last three decades, scholars have investigated the effectiveness of computer-assisted language learning (CALL). While CALL is gaining significance, the language skill that has received the least attention in research studies is writing (Rahimi et al., 2020). Previous research has investigated Iranian teachers' and students' perspectives on utilizing computers for language learning. Despite this, the influence of computer-assisted learning on language education lacks sufficient investigation (Fathali & Emadi, 2021; Rahimi et al., 2020).

Furthermore, previous studies have not dealt with the effect of modality on integrated writing performance among Iranian EFL learners. To the best of our knowledge, no study has ever examined Iranian EFL integrated writing in online and traditional modes; therefore, it will be illuminating to study integrated writing essays in the two modalities (online vs. paper-based) in the Iranian EFL writing assessment context.

Considering the above, the following research question was posed in this study:

-Is there any difference between EFL learners' integrated writing performance in online and paper-based modes?

Methodology

Design of the study

This study adopted a quantitative quasi-experimental design to investigate how integrated writing essays differ in online and paper-based modes. The students in two online and paper-based groups composed integrated essays. The essays were compared to examine if the mode of test delivery had any bearing on the integrated essays.

Context of the study

The current study was conducted in the English department at Persian Gulf University in Bushehr, Iran. According to the Iranian Ministry of Science, Research and Technology (MSRT), undergraduate academic writing is divided into two courses: advanced writing, which concentrates on paragraphs, and essay writing. The educational contexts in Iran do not provide sufficient instruction on writing paragraphs and essays to students. Therefore, this study provided a situation for teaching writing in an integrated approach to help learners learn the essentials of composing academic integrated essays in English based on the information obtained from a reading passage input. It took the researcher approximately three months to complete the study.

Participants

In this study, 44 male and female undergraduate students who majored in English language and literature at Persian Gulf University (PGU) participated. All the participants were native speakers of Persian and came from two intact classes. Their ages ranged between 20 and 25. Using convenience sampling, two essay writing classes were randomly assigned into two online and paper-based modes. In addition, two experienced writing instructors were recruited for this study. They were English professors at the PGU English department. Moreover, they were well-experienced in teaching/assessing writing.

According to the research ethical issues, the participants' privacy and confidentiality were respected, and they were asked to sign an informed consent form to participate. They were also assured that their data would be used only for research purposes. So, they agreed to contribute actively to the study period, which lasted for three months.

Instruments

To address the purpose of this study, three instruments were used. In the following, they are entirely elaborated.

Integrated writing test

The researcher presented a reading-to-write task to assess the students' ability to produce integrated writing in online and traditional groups. Each group requested the participants to share their thoughts on the desired topic.

Reading passage

A reading text and/or a listening prompt are needed to write an integrated writing essay. In this research, participants in both groups were given a two-page reading prompt about "capital punishment and its pros and cons" as a sample text to respond to a read-to-write (RW) task. Participants were asked to attentively peruse the given reading sample and accurately describe its topic, incorporating the significant details presented in the text. Noteworthy is that students who participated in online education were expected to use their online devices, like computers, to accomplish the assignment with the aid of Microsoft Word software. Furthermore, an online platform, LMS (learning management system), was created as virtual classrooms for students to participate in online meetings. Consequently, the online group's participants did the RW task via computers, while the face-to-face participants performed traditionally, using paper and pencil.

Integrated writing assessment rubric

For scoring participants' essays, a holistic rating rubric was used in this study that was developed by the TOEFL iBT test, which contains five rating domains within which different elements were taken into consideration when scoring the essays, which ranged from 0 (lowest) to 5 (highest) in terms of quality. The criteria included in this rubric are grounded on selecting appropriate and vital information from the passage, producing grammatically correct structures, and presenting a coherent and organized passage nearly connected to the main ideas of the source text.

Data collection procedure

Early in the study, a group of 44 students coming from two intact classes were randomly assigned to an online and a paper-based group. Before data collection, the students were informed about the study. They agreed to participate in the study voluntarily. They were also told that their data would be confidential and used only for research. The researchers also made several attempts to protect the participants' privacy and confidentiality and, in this way, maintain the ethicality of the study.

Throughout a 16-week of instruction (one session a week), the students in both groups were asked to write an essay in response to a reading prompt. The students in the paper-based group practiced integrated writing tasks using the conventional paper-and-pen procedure. The instructor provided corrective feedback on their integrated essays. In this way, they learned to negotiate different aspects of the integrated writing task with the teacher. On the other hand, students in the online group were taught using the LMS platform. Due to their vast experience with online teaching during the COVID-19 pandemic, the students were quite familiar with how to perform in the online class. The online students received similar instruction on integrated writing tasks as the paper-based group. Upon the treatment sessions, the students in the two groups were assessed using an integrated writing task. They were asked to write an argumentative essay using an accompanying reading passage. At the end of the study, 44 essays were gathered for later analysis.

Data analysis

This study used the TOEFL iBT integrated writing rubric to rate the students' texts in the two groups. The scale is a holistic one that analyzes the students' performance in terms of the appropriate selection of the information from the input, the presentation of the information in a coherent way, and the extent of usage and grammar errors. As the TOEFL iBT rating scale was originally developed for listen-to-write (LTW) writing tasks, Gebriel

(2009, 2010) modified the scale by omitting any reference to the listening component. This revised version was adopted in this study for rating the texts.

The essays were doubly rated by the researcher and an experienced rater. Each numerical score given to the integrated essays refers to the extent to which the appropriate information and grammatically correct structures were written.

In addition, using IBM SPSS statistics (version 22), an independent-sample t-test was run to compare the students' integrated writing performance in the online and paper-based groups.

Results

To address the research objective, which aimed to investigate the effect of test modality on the students' integrated writing performance, the students' integrated writing performances in the two groups (i.e., online and paper-based) were compared. The results of descriptive statistics are presented in Table 1 below.

Table 1. Descriptive statistics of learners' integrated writing in online and face-to-face modes

Integrated essays	Mode	N	Mean	Std. Deviation	Std. Error Mean
	Online	22	2.8182	1.33225	0.28404
Face-to-face	22	1.5909	0.66613	0.14202	

As Table 1 above shows, the online group ($M = 2.81$, $SD = 1.33$) had a higher mean score than the paper-based group ($M = 1.59$, $SD = 0.66$), indicating their better performance in producing integrated writing essays.

Table 2. Independent-sample T-tests for online and paper-based integrated writing tasks

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Integrated essays	10.334	0.003	3.865	42	0.000	1.22727	0.31756	0.58641	1.86814
Equal variances assumed									
Equal variances not assumed			3.865	30.882	0.001	1.22727	0.31756	0.57950	1.87505

An independent samples T-test was run, and the results, as displayed in Table 2, revealed that there was a significant difference between online ($M = 2.81$, $SD = 1.33$) and paper-based ($M = 1.59$, $SD = 0.66$; $t(42) = 3.86$, $p = 0.00$, two-tailed) groups. The magnitude of the differences in the means (mean difference = 9.27, 95% $CI = 0.58$ to 1.86) was significant, indicating that performances were significantly different in the two modalities.

Discussion

The analysis of the results suggested that the online mode of integrated writing test delivery significantly increased the quality of the integrated writing produced in the online group. In other words, utilizing the online modality noticeably enhanced the students' writing abilities. Overall, the findings of the present research agree with studies that underscore the potential of CBT for improving the performance of students (Dastpak et al., 2021; Guapacha Chamorro, 2022; MovahedFar et al., 2022; Marandi & Nami, 2012). Along the same line, one argument for the better performance of the students in the online group can be the learners' enhanced access to different language resources provided on the Internet. This finding agrees with what Rahimi et al. (2020) stated about being availed to the Internet. They claimed that students can easily access various educational materials through the Internet. This implies that finding information online is made faster and more accessible due to the multiple features embedded in the Web, such as spell and grammar checkers, online dictionaries, synonyms, etc.

Each feature can assist the online participants in paraphrasing the prompt text's main ideas and checking the grammatical structures' correctness when producing integrated writing essays. Since the TOEFL iBT test assesses students' proficiency in linking concepts and arranging data, online group students felt that employing word editing tools could facilitate the creation of an understandable and skillfully written essay. Using a word processing tool for online writing tasks assists in selecting synonyms and providing feedback on language mistakes, which may hinder comprehension. As Cochran-Smith (1991), Holland (1996), Wolfe et al. (1996), and Dastpak et al. (2021) mentioned, online editing and revising are much more accessible and more uncomplicated than revising in pencil and paper texts; hence, all the students are enthusiastic by using word processors for writing. Moreover, their attitudes toward writing improve when required to write using word processors. As a result, when composing integrated essays, the participants received assistance from word software, which aided them in selecting appropriate words and structuring their ideas effectively. Jin and Yan (2017) also reported that using a computer for writing leads to more grammatical, vocabulary-rich, and precise texts. Consequently, it implies that essays written using a computer could receive better grades than essays written manually.

On the other hand, the students in the paper-based group lacked adequate editing tools, resulting in their tendency to predominantly utilize copy-pasting techniques rather than producing logical and coherent essays. As a result, the quality of their essays did not meet the expected standards. The traditional group might have faced challenges when attempting to enhance the organization and coherence of their writing by adding or removing sentences, which could make the task difficult and time-consuming.

From another perspective, the better performance of the online group can be attributed to the psycholinguistic factors that might impact the students' work on the writing task. Krashen (2008) believed that self-confidence, anxiety, and the motivation to acquire knowledge are crucial to achieving a second language. The students in the online group appeared to possess superior self-efficiency and exhibited less anxiety while writing. According to Gunge and Taylor (1989), insufficiently produced writing texts suggest a person feels anxious about writing. In contrast, students who employed an online computer program to write their essays had longer pieces than those who used a pencil and paper in a group setting. This conclusion aligns with the results obtained by Russell and Haney (1997) and Jin and Yan (2017) who found that students who employed computers for their essays could generate nearly two times the quantity of writing compared to those who favored paper.

Similarly, Avazaghaei et al. (2022) found that integrated writing tasks in traditional paper-based contexts induced higher writing anxiety levels than in online mode. Additionally, online students exhibited a greater tendency to arrange their thoughts in multiple paragraphs in the online computer-assisted way due to their higher motivation than in the traditional mode. Furthermore, individuals who completed their tasks from home might have experienced reduced anxiety and enhanced motivation compared to students in a conventional classroom setting.

Moreover, from a textual standpoint, during the composing process of integrated writing, the content of the source texts can significantly influence the level of comprehension and the quality of the essays produced by test-takers. One aspect of the integrated writing tasks evaluation is comprehending the original information in the reading input (Cumming et al., 2005). It is conceivable that online sources were used to translate or rephrase some complex sentences that assisted participants in comprehending the text more effectively and generating a significant and precise rendition while keeping vital details intact. The idiosyncratic composing process in the integrated writing tasks called discourse synthesis is a demanding phase in the writers' attempt to judiciously select ideas from the source reading text and include them in well-developed and grammatically sound sentences (Plakans, 2009). Any component that can ease this complicated and burdensome process would aid learners in writing better essays. The present study's findings showed that online context, with its many provisions compared to the conventional paper-based testing mode, can reduce the composing burden on the learners in integrated writing. Hence, they could write better in the online mode.

Conclusion

In sum, the findings of the present study concluded that online modality positively affected the participants writing in online mode and can reduce psychological barriers such as anxiety and lack of self-efficiency and, as a result, emphasizes the supportive impact of using computer-assisted writing in the Iranian EFL learning and assessment contexts. Overall, the better performance of the participants in the online group than the traditional group provides several implications for the theory and practice of EFL integrated writing assessment.

The major theoretical implication of this study is its contribution to understanding how online education and using computers can impact the participants in L2 writing assessment in general and integrated writing assessment in particular.

As time passes, the online environment has developed new technologies that have led language teachers to consider online platforms as valuable resources to educate learners and enhance their learning achievements (Jiang et al., 2023). In addition, the findings of this study recommend several pedagogical implications for EFL-integrated writing assessment practice. Firstly, incorporating integrated writing exercises can significantly assist language educators in discovering effective teaching methods, particularly for students getting ready to take the TOEFL or IELTS exams. Secondly, teachers can effectively pinpoint students' writing weaknesses and develop strategies to overcome them by comprehending the areas where students struggle and encounter difficulties. This enables teachers to modify their teaching styles accordingly.

Furthermore, this study explored the advantages teachers can gain from utilizing Microsoft Word software in virtual educational settings while assessing students' writing skills. Achievements obtained by the online mode in this research can be implied in institutions investing in organizing online courses, specifically, online integrated writing classes. Last, the research findings can help organizations that administer high-stakes exams like PTE, TOEFL, and IELTS, which increasingly use integrated tasks in their online administration of the tests. Last, this research aligns with the conclusions of Najmi (2015), who identified that incorporating technology in education aids in fostering student engagement and can revise their writing with prompts and hints. Thus, it is necessary to include technology-driven (e.g., computer) assessments to provide a more comprehensive evaluation of students' learning. Assessing one's writing skills through computer software is crucial in enhancing writing abilities and achieving optimal outputs.

The study also faced some limitations. In terms of the design of the study, the prevalence of the COVID-19 pandemic made it challenging to arrange a physical gathering involving a large cohort of students. Hence, it was attempted to establish a classroom environment that could accommodate many individuals for the research. Furthermore, the poor internet connectivity in Iran caused a lack of optimal performance from the participants in the online group. Despite certain limitations, the researcher believes that this study offers new insights for scientific investigation into the impact of online and face-to-face modalities on the performance of EFL learners in integrated writing assessments. In this way, it contributes a novel aspect to the prevailing literature on integrated writing assessments in online learning contexts.

This research has thrown up several questions in need of further investigation. Further investigation is needed to determine the effectiveness of Microsoft Word in online education. To fully comprehend the superiority of Microsoft Word, it is recommended to compare it with other computer programs. Microsoft Word assisted the online group. Additional investigation is required to develop platforms that deliver the same experience as traditional pen and paper to minimize the possibility of corrective feedback in online media.

Furthermore, an online platform was employed to write integrated writing texts during the research. However, the online participants encountered problems related to their Internet connectivity. To compensate for the issues associated with these platforms, it is advisable to choose platforms that do not require high-speed Internet.

Since a convenience sampling procedure was used to conduct this research, the better performance of the online group in comparison to the traditional group can be attributed to various factors, including the participants' prior experience with the language (e.g., L2 background) and decreased anxiety levels. Consequently, further research is needed to

homogenize the two classes before the study. Furthermore, considering the project's design, the current study used a quasi-experimental design. However, future studies can use qualitative designs to study and understand topics from different perspectives. For example, future researchers can design semi-structured interviews to investigate Iranian EFL learners' perceptions when writing while using online education. Last, the study was conducted with two intact groups, each consisting of 22 students, limiting the study's generalizability. Hence, future studies are needed to recruit a larger sample to improve the generalizability of the study.

References

- Abdolrezapour, P. (2019). Applying computer-mediated active learning intervention to improve L2 listening comprehension. *Applied Research on English Language*, 8(4), 511-530. <https://dx.doi.org/10.22108/are.2019.115355.1424>
- Abdolrezapour, P., & Ghanbari, N. (2022). Emotional-based pedagogy and facilitating EFL learners' perceived flow in online education. *Frontiers in Psychology*, 13, 960287. <https://doi.org/10.3389/fpsyg.2022.960287>
- Avazaghaei, M., Ghanbari, N., & Abbasi, A. (2022). Investigating the effect of writing task type (i.e., Integrated vs. Independent) on the anxiety level of the students: a mixed-methods study. Unpublished MA thesis. Persian Gulf University, Bushehr, Iran.
- Barkaoui, K. (2015). Test takers' writing activities during the TOEFL iBT® writing tasks: A stimulated recall study. *ETS Research Report Series*, 2015(1), 1-42. <http://dx.doi.org/10.1002/ets2.12050>
- Barkaoui, K. (2016). What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *The Modern Language Journal*, 100(1), 320-340. <https://doi.org/10.1111/modl.12316>
- Barkaoui, K., & Knouzi, I. (2018). The effects of writing mode and computer ability on L2 test-takers' essay characteristics and scores. *Assessing Writing*, 36, 19-31. <https://doi.org/10.1016/j.asw.2018.02.005>
- Blackhurst, A. N. D. R. E. W. (2005). Listening, reading and writing on computer-based and paper-based versions of IELTS. *Research Notes*, 21(1), 14-17.
- Braine, G. (1989). Writing in science and technology: An analysis of assignments from ten undergraduate courses. *English for Specific Purposes*, 8(1), 3-15. [https://doi.org/10.1016/0889-4906\(89\)90003-3](https://doi.org/10.1016/0889-4906(89)90003-3)
- Carson, J. (2001). A task analysis of reading and writing in academic contexts. In D. Belcher & A. Hirvela (Eds.). *Linking literacies: Perspectives on L2 reading-writing connections* (pp. 48-83). Ann Arbor, MI: University of Michigan Press.
- Chambers, L. (2008). Computer-based and paper-based writing assessment: A comparative text analysis. *Research Notes*, 34, 9-15.
- Chapelle, C. A., & Douglas, D. (2006). *Assessing language through computer technology*. Cambridge University Press.
- Choi, Y. D. (2022). Validity of score interpretations on an online English placement writing test, *Language Testing in Asia*, 12(1), 42. doi: [10.1186/s40468-022-00187-0](https://doi.org/10.1186/s40468-022-00187-0)
- Cochran-Smith, M. (1991). Word processing and writing in elementary classrooms: A critical review of related literature. *Review of educational research*, 61(1), 107-155. <https://doi.org/10.3102/00346543061001107>

- Cumming, A. (1997). Assessing writing. In C. Clapham, & D. Corson (Eds.), *Encyclopedia of language and education: Language testing and assessment* (pp. 51–63). Dordrecht, The Netherlands: Kluwer.
- Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing*, 10(1), 5-43. <https://doi.org/10.1016/j.asw.2005.02.001>
- Cumming, A., Kim, T. Y., & Eouanzoui, K. B. (2007). Motivations for ESL writing improvement in pre-university contexts. In Boscolo, P. and Hidi, S.(eds.) *Writing and Motivation*. Vol. 19 (pp. 93-111). Oxford: Elsevier
- Cumming, G. (2011). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York: Routledge. <https://doi.org/10.4324/9780203807002>
- Cumming, A., Grant, L., Mulcahy-Ernt, P., & Powers, D. E. (2004). A teacher-verification study of speaking and writing prototype tasks for a new TOEFL. *Language Testing*, 21(2), 107-145. <https://doi.org/10.1191/0265532204lt278oa>
- Cumming, A. (2013). Assessing integrated writing tasks: Promises and perils. *Language Assessment Quarterly*, 10(1), 1-8. <https://doi.org/10.1080/15434303.2011.622016>
- Dastpak, M., Riasati, M. J., & Hadipourfard, E. (2021). Investigating IELTS Paper Mode vs. Computer Mode: Evidence from Academic Writing Test. *Journal of Language and Translation*, 11(2), 175-193. <https://doi.org/10.30495/tlt.2021.682816>
- Daiute, C. (1986). Physical and cognitive factors in revising: Insights from studies with computers. *Research in the Teaching of English*, 20(2), 141-159. <https://www.jstor.org/stable/40171074>
- Delaney, Y. A. (2008). Investigating the reading-to-write construct. *Journal of English for academic purposes*, 7(3), 140-150. <https://doi.org/10.1016/j.jeap.2008.04.001>
- wiyanti, K. E., & Suwastini, N. K. A. (2021). Assessment for writing skills in online learning. *Lingua Scientia*, 28(1), 8-19. <https://doi.org/10.23887/ls.v28i1.29069>
- Elmahdi, I., Al-Hattami, A., & Fawzi, H. (2018). Using Technology for Formative Assessment to Improve Students' Learning. *Turkish Online Journal of Educational Technology-TOJET*, 17(2), 182-188. <https://eric.ed.gov/?id=EJ1176157>
- Esmaili, H. (2002). Integrated reading and writing tasks and ESL students' reading and writing performance in an English language test. *Canadian Modern Language Review*, 58(4), 599-620. <https://doi.org/10.3138/cmlr.58.4.599>
- Fathali, S., & Emadi, A. (2021). CALL research in Iran: An integrative review of the studies between 2007 and 2019. *Computer-Assisted Language Learning Electronic Journal*, 22(3), 33-51. <http://callej.org/journal/22-3/Fathali-Emadi2021.pdf>
- Feak, C., & Dobson, B. (1996). Building on the impromptu: a source-based academic writing assessment. *College ESL*, 6(1), 73-84. <https://eric.ed.gov/?id=EJ530884>
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College composition and communication*, 32(4), 365-387. <https://doi.org/10.2307/356600>
- Gebril, A. (2009). Score generalizability of academic writing tasks: Does one test method fit it all?. *Language Testing*, 26(4), 507-531. <https://doi.org/10.1177/0265532209340188>
- Gebril, A. (2010). Bringing reading-to-write and writing-only assessment tasks together: A generalizability analysis. *Assessing writing*, 15(2), 100-117. <https://doi.org/10.1016/j.asw.2010.05.002>

- Ghanbari, N., & Abdolrezapour, P. (2021). Using emotional intelligence in an EFL integrated writing assessment. *Studies in Educational Evaluation*, 70, 101017. <https://doi.org/10.1016/j.stueduc.2021.101017>
- Ghanbari, N., & Nowroozi, S. (2021). The practice of online assessment in an EFL context amidst COVID-19 pandemic: views from teachers. *Language Testing in Asia*, 11, 1-18. <https://doi.org/10.1186/s40468-021-00143-4>
- Gholami, J., & Alinasab, M. (2017). Source-based tasks in writing independent and integrated essays. *International Journal of Instruction*, 10(3), 127-142. <https://doi.org/10.12973/iji.2017.1039a>
- Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & education*, 57(4), 2333-2351. <https://doi.org/10.1016/j.compedu.2011.06.004>
- Green, T. O. N. Y., & Maycock, L. O. U. I. S. E. (2004). Computer-based IELTS and paper-based versions of IELTS. *Research Notes*, 18, 3-6.
- Guapacha Chamorro, M. E., (2022). Cognitive validity evidence of computer- and paper-based writing tests and differences in the impact on EFL test-takers in classroom assessment, *Assessing writing*, 51, 100594, <https://doi.org/10.1016/j.asw.2021.100594>.
- Gunge, B. W., & Taylor, V. (1989). Writing Apprehension in Second Language Writers. In D. Johnson and D. Roen (Ed). *Richness in Writing*: 235-248.
- Hafner, C. A., & Miller, L. (2011). Fostering learner autonomy in English for science: A collaborative digital video project in a technological learning environment. *Language, Learning and Technology*, 15(3), 68-86. <http://llt.msu.edu/issues/october2011/hafnermiller.pdf>
- Hirvela, A. (2016). *Connecting reading & writing in second language writing instruction (2nd ed.)*. Ann Arbor: University of Michigan Press.
- Holland, J. L. (1996). Exploring careers with a typology: What we have learned and some new directions. *American psychologist*, 51(4), 397-406. <https://doi.org/10.1037/0003-066X.51.4.397>
- Horkay, N., Bennett, R. E., Allen, N., Kaplan, B., & Yan, F. (2006). Does it matter if I take my writing test on computer? An empirical study of mode effects in NAEP. *Journal of Technology, Learning, and Assessment*, 5(2), n2. <https://ejournals.bc.edu/index.php/jtla/article/view/1641>
- Horowitz, D. M. (1986). What professors actually require: Academic tasks for the ESL classroom. *TESOL quarterly*, 20(3), 445-462. <https://doi.org/10.2307/3586294>
- Huang, H. T. D., & Hung, S. T. A. (2010). Examining the practice of a reading-to-speak test task: Anxiety and experience of EFL students. *Asia Pacific Education Review*, 11, 235-242. <https://doi.org/10.1007/s12564-010-9072-6>
- Jin, Y., & Yan, M. (2017). Computer literacy and the construct validity of a high-stakes computer-based writing assessment. *Language Assessment Quarterly*, 14(2), 101-119. <https://doi.org/10.1080/15434303.2016.1261293>
- Jiang, P., Namaziandost, E., Azizi, Z., & Razmi, M. H. (2023). Exploring the effects of online learning on EFL learners' motivation, anxiety, and attitudes during the COVID-19 pandemic: a focus on Iran. *Current Psychology*, 42(3), 2310-2324. <https://doi.org/10.1007/s12144-022-04013-x>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge university press.

- Knoch, U., & Sitajalabhorn, W. (2013). A closer look at integrated writing tasks: Towards a more focused definition for assessment purposes. *Assessing writing*, 18(4), 300-308. <https://doi.org/10.1016/j.asw.2013.09.003>
- Krashen, S. (2008). 4. The Comprehension Hypothesis Extended. In T. Piske & M. Young-Scholten (Ed.), *Input Matters in SLA* (pp. 81-94). Bristol, Blue Ridge Summit: Multilingual Matters. <https://doi.org/10.21832/9781847691118-007>
- Lee, Y. J. (2002). A comparison of composing processes and written products in timed-essay tests across paper-and-pencil and computer modes. *Assessing Writing*, 8(2), 135-157. [https://doi.org/10.1016/S1075-2935\(03\)00003-5](https://doi.org/10.1016/S1075-2935(03)00003-5)
- Lee, Y. W., Breland, H., & Muraki, E. (2004). Comparability of TOEFL CBT writing prompts for different native language groups. *ETS Research Report Series*, 2004(1), i-56. <https://doi.org/10.1002/j.2333-8504.2004.tb01951.x>
- Leki, I., & Carson, J. (1997). "Completely different worlds": EAP and the writing experiences of ESL students in university courses. *TESOL quarterly*, 31(1), 39-69. <https://doi.org/10.2307/3587974>
- Lewkowicz, J. A. (1997). The integrated testing of a second language. *Encyclopedia of language and education*, 7, 121-130. https://doi.org/10.1007/978-1-4020-4489-2_12
- Li, J. (2006). The mediation of technology in ESL writing and its implications for writing assessment. *Assessing Writing*, 11(1), 5-21. <https://doi.org/10.1016/j.asw.2005.09.001>
- Liu, Y., & Liu, J. (2023). Comparing computer-based and paper-based rating modes in an English writing test. *Assessing Writing*, 57, 100771. <https://doi.org/10.1016/j.asw.2023.100771>
- Lumley, T. (2005). *Assessing second language writing: The rater's perspective* (Vol. 3). Frankfurt: Peter Lang.
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- Marandi, S. S. (2002). Teaching English in the new millennium: CALL in Iran. In A. A. Rezaei (Ed.), *Proceedings of the First Conference on Issues in English Language Teaching in Iran* (pp. 205-220). Tehran, Iran: University of Tehran Press.
- Marandi, S. S., & Nami, F. (2012). Web-Based writing lessons in EFL contexts: Instruction on coherent writing. *Journal of Studies in Learning and Teaching English*, 1(2), 105-136. https://jslte.shiraz.iau.ir/article_518889.html
- Miller, M. D., & Legg, S. M. (1993). Alternative Assessment in a HighStakes Environment. *Educational Measurement: Issues and Practice*, 12(2), 9-15. <https://doi.org/10.1111/j.1745-3992.1993.tb00528.x>
- Mohamadi, Z. (2018). Comparative effect of online summative and formative assessment on EFL student writing ability. *Studies in Educational Evaluation*, 59, 29-40. <https://doi.org/10.1016/j.stueduc.2018.02.003>
- Mompean, A. R. (2010). The development of meaningful interactions on a blog used for the learning of English as a Foreign Language. *ReCALL*, 22(3), 376-395. <https://doi.org/10.1017/S0958344010000200>
- MovahedFar, S., Abbasian, G.R., Ameri, A. R. (2022). Computer-based dynamic assessment of EFL learners' writing performance: Evidence from both cognitive and emotive domains. *Journal of Language and Translation*, 12 (3), 1-19. https://ttl.stb.iau.ir/article_692048_a72615738a2ca5e74f3ac35c2f04ade8.pdf

- Najmi, K. (2015). The effect of genre-based approach on enhancing writing skill of Iranian law students. *Modern Journal of Language Teaching Methods*, 5(2), 474. https://mjltm.org/browse.php?a_id=52&sid=1&slc_lang=en&ftxt=1#page=474
- Plakans, L. (2008). Comparing composing processes in writing-only and reading-to-write test tasks. *Assessing Writing*, 13(2), 111-129. <https://doi.org/10.1016/j.asw.2008.07.001>
- Plakans, L. (2009a). Discourse synthesis in integrated second language writing assessment. *Language Testing*, 26(4), 561-587. <https://doi.org/10.1177/0265532209340192>
- Plakans, L. (2009b). The role of reading strategies in integrated L2 writing tasks. *Journal of English for academic purposes*, 8(4), 252-266. <https://doi.org/10.1016/j.jeap.2009.05.001>
- Plakans, L. (2010). Independent vs. integrated writing tasks: A comparison of task representation. *TESOL Quarterly*, 44(1), 185-194. <https://www.jstor.org/stable/27785076>
- Plakans, L., & Gebril, A. (2012). A close investigation into source use in integrated second language writing tasks. *Assessing Writing*, 17(1), 18-34. <https://doi.org/10.1016/j.asw.2011.09.002>
- Plakans, L., Gebril, A., & Bilki, Z. (2019). Shaping a score: Complexity, accuracy, and fluency in integrated writing performances. *Language Testing*, 36(2), 161-179. <https://doi.org/10.1177/0265532216669537>
- Rahimi, A., Jahangard, A., & Norouzizadeh, M. (2020). Students' attitudes towards computer-assisted language learning and its effect on their EFL writing. *International Journal of Learning and Teaching*, 12(3), 144-152. <https://doi.org/10.18844/ijlt.v12i3.4767>
- Raimes, A. (1998). *Exploring through writing: A process approach to ESL composition*. Cambridge University Press.
- Risemberg, R. (1996). Reading to write: Self-regulated learning strategies when writing essays from sources. *Literacy Research and Instruction*, 35(4), 365-383. <https://doi.org/10.1080/19388079609558221>
- Rose, S. K., & Weiser, I. (2010). Going public: What writing programs learn from engagement.
- Russell, M., & Haney, W. (1997). Testing writing on computers. *Education policy analysis archives*, 5, 3-3. <https://eric.ed.gov/?id=EJ580763>
- Soltani, A., & Kheirzadeh, S. (2017). Exploring EFL students' use of writing strategies and their attitudes towards reading-to-write and writing-only tasks. *Journal of Language and Linguistic Studies*, 13(2), 535-560.
- Webber, K. L. (2012). The use of learner-centered assessment in US colleges and universities. *Research in Higher Education*, 53(2), 201-228. <https://doi.org/10.1007/s11162-011-9245-0>.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.
- Weigle, S. C. (2004). Integrating reading and writing in a competency test for non-native speakers of English. *Assessing writing*, 9(1), 27-55. <https://doi.org/10.1016/j.asw.2004.01.002>.
- Weir, C., Yan, J., O'Sullivan, B., & Bax, S. (2007). Does the computer make a difference?: The reaction of candidates to a computer-based versus a traditional handwritten form of the IELTS Writing component: effects and impact.

- In *International English Language Testing System (IELTS) Research Reports 2007: Volume 7* (pp. 1-37). [Canberra]: IELTS Australia and British Council.
- Whithaus, C., Harrison, S. B., & Midyette, J. (2008). Keyboarding compared with handwriting on a high-stakes writing assessment: Student choice of composing medium, raters' perceptions, and text quality. *Assessing Writing*, 13(1), 4-25. <https://doi.org/10.1016/j.asw.2008.03.001>
- Wolfe, E. W., Bolton, S., Feltovich, B., & Niday, D. M. (1996). The influence of student experience with word processors on the quality of essays written for a direct writing assessment. *Assessing writing*, 3(2), 123-147. [https://doi.org/10.1016/s1075-2935\(96\)90010-0](https://doi.org/10.1016/s1075-2935(96)90010-0)
- Wolfe, E. W., & Manalo, J. R. (2005). An investigation of the impact of composition medium on the quality of scores from the TOEFL writing section: A report from the broad-based study (TOEFL Research Report. No. RR-72). *Princeton, NJ: Educational Testing Service*.
- Yang, H. C. (2009). *Exploring the complexity of second language writers' strategy use and performance on an integrated writing test through structural equation modeling and qualitative approaches*. The University of Texas at Austin.
- Yu, G. (2013). From integrative to integrated language assessment: Are we there yet?. *Language Assessment Quarterly*, 10(1), 110-114. <https://doi.org/10.1080/15434303.2013.766744>.
- Yu, W., & Iwashita, N. (2021). Comparison of test performance on paper-based testing (PBT) and computer-based testing (CBT) by English-majored undergraduate students in China. *Language Testing in Asia* 11, 32. <https://doi.org/10.1186/s40468-021-00147-0>.
- Zlatović, M., Balaban, I., & Kermek, D. (2015). Using online assessments to stimulate learning strategies and achievement of learning goals. *Computers & education*, 91, 32-45. <https://doi.org/10.1016/j.compedu.2015.09.012>