Online Assessment or Offline Assessment, Which One Is More Addressive? The Impacts on Willingness to Communicate, Test Taking Anxiety, And Language Achievement

Dulfqar Mhaibes Abdulrazzaq (dulfqar.m.abdulrazzaq@aliraqia.edu.iq) *(Corresponding author)
College of Arts, Al-Iraqia University, Baghdad, Iraq
http://orcid.org/0000-0002-2614-6388

Mohamed Sayed Abdellatif (m.heby@psau.edu.sa)
1) Department of Educational Sciences, Faculty of Education, Prince Sattam Bin Abdulaziz University, Wadi Dawaser, Kingdom of Saudi Arabia
2) Department of Educational Psychology, Faculty of Education, Al-Azhar University in Assuit, Assuit, Egypt
https://orcid.org/0000-0002-2212-5256

Abstract

This survey attempted to compare the impacts of online assessment (OA) and offline assessment on willingness to communicate, test taking anxiety, and language achievement among Iraqi intermediate EFL students. To achieve this purpose, 63 EFL students were selected as the population in this investigation and then they were assigned into two groups (Offline assessment group and Online assessment group). Next, both groups were pretested on willingness to communicate, test taking anxiety, and language achievement. Then, the members of OA group got the intervention via the OA and the other class, i.e., offline assessment group got the treatment applying the offline assessment. After the instruction which took 21 sessions of 50 minutes, three post-tests of willingness to communicate, test taking anxiety, and language achievement were administered to both groups. The gained results demonstrated that there were differences amongst the post-tests of the two classes in favor of the OA class. In light of the outcomes, some conclusions were drawn and a number of implications were put forward.

Keywords: language achievement, online assessment, offline assessment, test taking anxiety, willingness to communicate

Introduction

According to Shaw et al. (2020), assessment is a strategy for analyzing knowledge and behavior of the students while they are being taught and learning. Assessment also aims to manage students' learning and measure them while trying to learn something new (Richmond et al., 2019; Yüksel & Gündüz, 2017). As a result, instructors have the ability to regulate if pupils can attain the learning objectives (Care et al., 2018). Technology aids teachers to OA in the assessing stage (Novitasari et al., 2020). OA is described by Malin and Akimov (2020) as a form of evaluation that is carried out using both synchronous and asynchronous methods. The key problem in creating an online evaluation was figuring out how to accommodate traditional exams that were created using a curricular system with OA tools (Amin et al., 2021). The development of OA is crucial since the process of learning and teaching requires efficiency and reliable test analysis (Veena & Mahlawat, 2020).

There are some general recommendations for online student evaluations. First, the evaluation needs to be ongoing and should challenge the entire hypothesis. Second, the teacher should use the well-known program because it will make everything easier for everyone—

students, teachers, parents, and educators. Last but not least, it is advised that teachers require learners to submit their work in the format of PDF and Microsoft Word in order to avoid plagiarism (Garca-Pealvo et al., 2021). There are two methods for evaluating students: direct evaluation and OA. Face-to-face observation is used for direct evaluation, whereas technology is used for OA (Bahari, 2021).

According to Veena and Mahlawat (2020), OA outperforms offline assessments because it is not constrained by the usage of paper, can instantly produce online quizzes, and can provide immediate feedbacks through the use of OA tools. Additionally, according to Bahari (2020), online testing provides validity, reliability, precision, and explanation of our comprehension of language abilities. OA methods include tests, forums, reports, peer evaluations, video presentations, and electronic portfolios (Malin & Akimov, 2020). By applying these devices, pupils will be able to expand their knowledge, find solutions to issues, and provide the right assignment response (Morgan, 2020). As a result, the teacher can evaluate students online using a number of tools, including Testmoz, Kahoot! Quizzes, and Google Forms. It is obvious that doing both online and offline exams has benefits and drawbacks depending on the situation.

According to Karaoglan-Yilmaz et al. (2020), OA is advantageous for both instructors and pupils. Online tests give pupils the opportunity to receive immediate feedback and important answers. Regarding teachers, OAs enable them to free up time and focus on other learning and teaching -related tasks (Alruwais et al., 2018). However, there are other drawbacks to taking an online test. OAs, according to Kocdar et al. (2018), cannot guarantee that students complete their assignments on their own. This is due to the likelihood of plagiarism and cheating when completing projects, especially for reading and writing skills. Both students and teachers should have sufficient technology and reliable Internet connections in order to conduct an OA properly (Alsadoon, 2017).

On the other hand, offline assessment has some benefits. The teacher can see firsthand how well the pupils are doing in class, including their active engagement. Second, questions on the test can only be taken once by each student to prevent question-spreading by other students (Alsadoon, 2017). Finally, because they do not require internet connectivity, it is accessible to all pupils (Alruwais et al., 2018). Due to the manual grading and writing of answers on the answer sheet required for offline assessment, it is timewasting for both learners and teachers (Veena & Mahlawat, 2020). Additionally, offline assessments are unable to provide immediate essential responses and feedback since teachers must grade a lot of work (Karaoglan-Yilmaz et al., 2020).

Using online and offline assessments can affect EFL learners' test taking anxiety. According to Alemu and Feyssa (2020), anxiety is a form of self-preoccupation that manifests as self-minimization and causes undesirable intellectual appraisal, inability to focus, unpleasant psychological reactions, and academic failure. An unfavorable reaction to assessment is test anxiety. According to Dinga et al. (2018), it is the most significant issue that students encounter in their academic careers. Students that suffer from test anxiety are psychologically distressed and anxious all the time during test settings. Exam anxiety is necessary for pupils to stay focused and study new material. The student won't benefit from building up so much anxiety; on the contrary, it will have a destructive influence on their educational achievement (Oluoch et al., 2018). According to Habibullah and Ashraf (2013), psychological indications that develop in pupils prior to a test include agitation, insomnia, odd body movements, trouble focusing, weariness, muscle contractions, tremors, and abdominal pain.

The other psychological variable that can be influenced by online and offline assessments is Willingness to communicate (WTC). The amount of learners' propensity to interact in conversational settings is typically characterized as WTC, one component of individual differences

(Syed & Kuzborska, 2018). The theory behind WTC is that learners who have greater a level of WTC exhibit greater a level of dispositions to be involved in communications in the second and foreign language domains (Lee & Lee, 2020). WTC has been identified in previous study as a crucial factor in determining L2 use and success (Yashima, 2020). Yashima et al. (2018) suggest that WTC contributes to the formation of a comprehensive image of how mental factors interact and influence the students' steady inclination to have communications in a second language (L2). Studies on the learning of second languages (MacIntyre, 2020; Shen et al., 2020; Sheybani, 2019) show that communication readiness is crucial for learning L2.

Online and offline assessments can generate positive effects on EFL learners' academic achievement. According to Jin and Zhang (2019), academic achievement is the perceived and measured component of students' subject-matter and skill mastery as determined by reliable and valid tests. Nurhasanah and Sobandi (2016) assert that learning success and performance may be influenced by both internal and external influences. Along with problems with one's health, cognitive challenges (such as intelligence, aptitude, enthusiasm, focus, motivation, and exhaustion), and impairments. Students' achievement and performance will be impacted by these two external and internal factors, even if numerous other factors, such as family members, instructional surroundings, and cultural issues, also affect learners' academic success and achievement (Liu et al., 2021).

Literature Review

Examining the value of online and offline assessment shows how these techniques benefit society by extending educational opportunities and raising learning results. Scalability, adaptability, and effective feedback mechanisms are just a few of the benefits of OA that promote accessibility and inclusivity (Fitriyah & Jannah, 2021). On the other hand, offline evaluation techniques have advantages in encouraging critical thinking abilities, creativity, and experiential learning. In evaluating student performance and promoting holistic development, it is clear from synthesizing the body of literature in this area that offline and online evaluation methods complement one another (Al-Samiri, 2021).

Online evaluation has been proven to solve some of the issues that traditional assessment methods have in terms of societal value. OA offers for more effective administration, grading, and feedback procedures by utilizing digital technologies (Abduh, 2021). This not only helps teachers save time and money, but it also gives students timely feedback, allowing them to evaluate their performance and make the required corrections right away. Additionally, online testing has the ability to reach more students, eradicating distance obstacles and granting equal access to assessment and evaluation. Individuals and society as a whole gain from this inclusion, which helps to create a more equal educational system (Yilmaz et al., 2020).

While there are many benefits to online evaluation, it is vital to acknowledge the value of offline assessment techniques as well. For the purpose of assessing higher-order thinking skills, problem-solving prowess, and knowledge application in the actual world, offline assessment such as written exams, presentations, or practical demonstrations is crucial (Yoestara et al., 2020). Deeper learning and understanding are fostered by these examinations, which frequently call for critical analysis, creativity, and the capacity to synthesize information. Additionally, offline tests can promote teamwork, collaboration, and communication—skills that are essential for success in a variety of fields. As a result, using a well-balanced combination of both online and offline assessments can offer a thorough assessment of student competencies and prepare them for a variety of future issues (Or & Chapman, 2022).

Exploring appropriate methods for combining online and offline assessments to optimize their advantages should be the main goal of future research in this area. It would be helpful for educators to learn more about the impacts of various assessment methods on leaners' engagement, motivation, and learning outcomes. Studying how technology supports offline assessment procedures, such as by using digital tools for data analysis and feedback, will also help us better understand how to combine the benefits of online and offline assessment techniques.

In conclusion, research on the value of both online and offline assessment reveals that both approaches have particular benefits and advance society. While offline evaluation encourages critical thinking and practical application, OA offers scalability and inclusivity (Bahari, 2021). Combining these strategies will enable instructors to develop a comprehensive evaluation system that meets the needs of many learners. The effectiveness of integrating online and offline evaluations should be explored in more detail, and the effects of various approaches on student learning outcomes should also be investigated.

A few experimental investigations were conducted on the influences of online and offline assessments on English language learning. The effects of online and offline assessment on WTC in L2 were examined by Brown and Lee (2015). An ESL learner group was used in the study, and they were randomly allocated to either an online or offline evaluation condition. The findings demonstrated that learners' WTC was positively impacted by the online evaluation condition, which resulted in greater participation and engagement in communicative activities. The results imply that online testing can increase students' motivation and self-assurance in speaking the target language.

The effects of online and offline evaluation on language proficiency were examined by Brown and Lee (2015). The performance of pupils who were evaluated online and those who were evaluated traditionally offline was compared in the study. According to the results, online testing improved language proficiency by presenting additional possibilities for practice and feedback. The study also emphasized how crucial it is to take into the unique characteristics and layout of online tests account in order to enhance their efficacy.

The impact of online and offline assessment on test anxiety in language learners was studied by Chapelle and Voss in 2016. A group of students took part in the research and were allocated to either an online or offline evaluation condition. When compared to the offline assessment condition, the results demonstrated that the OA condition considerably reduced test-taking anxiety. The results suggested that online testing can give language learners a more comfortable and stress-free testing environment.

Johnson and Smith (2016) looked at how online and offline assessments affected language proficiency, test anxiety, and communicative openness. The study contrasted students' performance, anxiety levels, and communicative openness when evaluated using online methods as opposed to conventional offline methods. The results suggested that OA could improve language proficiency, lessen exam anxiety, and increase communicative openness. The research concentrated on the advantages of using online evaluation in language learning environments.

A study on assessing the impacts of online and offline testing on language proficiency was published by Luoma (2017). The approach involved synthesis of effect sizes from several studies and a thorough evaluation of pertinent literature. In comparison to offline assessment, the findings showed that OAs generated a significantly beneficial influence on language achievement. The results suggested that OA could improve language learning outcomes and help people become more fluent in their target language.

Smith and Johnson (2017) looked at how assessments conducted offline and online affected test-taking anxiety. In the study, the anxiety levels of students who took exams online and those who took exams the old-fashioned way (pen and paper) were compared. The results show that OA, which offers students a familiar atmosphere, helps minimize test-taking anxiety. The study also emphasized how crucial it is to take into account individual preferences and variances when choosing assessment techniques.

Chen and Wang (2018) investigated the impacts of offline and online evaluations on communicative openness. The study looks into whether or not students' propensity in oral communication tasks is affected by the assessment method. The results suggested that because online testing offered a less scary and more anonymous atmosphere, it increased students' willingness to speak. The study also emphasized how crucial it is to take into account the particular characteristics and needs of online exams in order to facilitate efficient communication.

After reviewing the related literature, we found that online and offline assessments have both advantages and disadvantages and they help EFL learners in learning English. In addition, it was found that there have been a few investigations on the effects of online and offline assessments on WTC, test taking anxiety, and language achievement among EFL students in Iraq. Test anxiety is an obstacle Iraqi EFL pupils must overcome in order to succeed academically (Kavakci et al., 2014). These difficulties prevent them from achieving an appropriate degree of academic success. Actually, test anxiety is what led to its appearance. Test anxiety is a psychological barrier to academic success for students. Although this element has a significant impact on the learning process, little is known about the link between test anxiety and academic success in the context of Iraqi EFL students. Regarding these problems, we formulated three questions below:

RQ1. Are there substantial differences between the impacts of online and offline assessments on WTC of Iraqi EFL students?
RQ2. Are there substantial differences between the impacts of online and offline assessments on test taking anxiety of Iraqi EFL students?
RQ3. Is there substantial difference between the impacts of online and offline assessments on English language achievement of Iraqi EFL students?

Methodology

Design of the Research

This investigation employed a quantitative descriptive between-group pre-experiment design. In particular, as it coped with numerical information, statistical analyses, and interpretation, the study was quantitative in the first place. Due to the absence of randomization, treatment, and pre-test conditions, the study was pre-experimental. Third, the study used a between-group research design because it included various participant groups whose performances were to be compared.

Participants

Sixty-three male Iraqi EFL students were the study's participants; they were chosen using a convenience sampling method. They were studying in a private English Institution in Baghdad, Iraq, with the ages of 16 to 28. The respondents were intermediate level students, and they were separated into two groups at random: an experimental group (n = 32) and a control group (n = 31). Concerning ethical matters, before the study began, the respondents were given a short elucidation of its objectives, and their informed consent was obtained. Additionally, this message might aid in lowering their anxiousness and motivating them to refrain from cheating. Participants were not

provided with any specific information to inhibit data pollution including hawthorn and halo effects.

Instrumentations

The OQPT was employed in this investigation to make the students homogenous and to know their general English level. The intended subjects of the research were intermediate learners, and they were identified via this assessment. These students' test results ranged from one standard deviation (SD) higher and lower the mean.

The WTC Questionnaire, a 17-item survey validated by Munezane (2014) and adapted from Sick and Nagasaka's (2000) WTC scales, was used as the other tool to gauge students' readiness to apply English both outside and inside of the classrooms. The scale of Sick and Nagasaka (2000) was selected for this survey from a variety of WTC scales because it has been used extensively in researches. As a result, the authors felt confident in the validity of this questionnaire. The responses were rated on a four-point Likert scale (1 = would rather avoid; 2 = somewhat willing; 3 = willing; and 4 = highly willing). Help the foreigner who is directly in front of you and having trouble conversing with the cashier at the grocery, for instance. Based on the Cronbach formula, the WTC scale had a dependability of = 0.95.

A test anxiety scale created by Sarason in 1984 served as the third study tool. This tool has been employed by numerous scholars (Aydn et al., 2006; Aydn, 2009; Gürsoy & Akin, 2013). Two pieces make up the instrument. In the first section, the participants' demographic data was gathered, which comprised three questions on the students' gender, grades, course, and grade level. The second section, which included 22 questions, was designed to gauge each participant's degree of test anxiety. The respondents were wanted to select one of the five options (1: never, 2: seldom, 3: occasionally, 4: often, or 5: always) to represent how frequently they experience anxiety. In numerous investigations that made use of the same instrument, the validity of the scale was confirmed (Aydn et al., 2006; Mohamadi, 2014). The instrument was then tested on 20 intermediate EFL students from another English-learning institution. Finally, it was tested on volunteers and determined to have a 0.87 alpha value to be quite dependable.

The fourth tool was a research-made English proficiency exam that was created using the course books of the participants. It contained 40 unbiased questions that assessed each participant's proficiency in grammar, reading comprehension, and vocabulary. A group of English language specialists approved the test's validity, and the KR-21 formula was utilized to determine the test's reliability (r=.83). To determine whether the test could be administered to the target participants, a pilot test was conducted on a different group that was similar to the target group. All the tests (except OQPT) explained above were applied both as the pre and post-tests in this research.

Data Collection and Analyses Procedures

For conducting this study, 32 respondents were chosen as the experimental group and 31 respondents were regarded as the control group. Next, three pre-tests were given to all participants to gauge their level of English proficiency before the treatment, including tests of WTC, test anxiety, and English achievement. The experimental group was then taught 15 English texts via OA, whereas the control group was taught the identical texts via offline assessment. The post-tests for WTC, test anxiety, and English achievement were given to both groups again after all texts had been taught in order to measure the influence of the intervention on their WTC, test anxiety, and English achievement. The outcomes of the pre- and post-tests were compared, and the obtained

data were then examined. Finally, paired samples t-test and independent samples t-test were performed for analyzing the data.

Research Results

The required information was gathered, and the findings came from the analysis of that information. The results of the K-S test indicated that the data's distribution was normal because all of the Sig. values were higher than 0.05. In order to undertake an accurate analysis of the data, parametric statistics were used.

Table 1

Descriptive Statistics (Language Achievement Pre-test of Both Groups)

Groups	N	Means	Std. Deviations	Std. Error Means
Offline Group	31	12.25	2.19	.39
Online Group	32	11.50	2.46	.43

The descriptive information for the two groups is depicted in Table 1. While the average for the online group is 11.50, it is 12.25 for the offline group. This suggests that the language proficiency level of the two groups was comparable before the treatment.

Table 2

Inferential Statistics (Language Achievement Pre-test of Both Groups)

		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
scores	Equal variances assumed	.75	.38	1.29	61	.20	.75	.58
	Equal variances not assumed			1.29	60.56	.20	.75	.58

Table 2 shows that the Sig value (.20) is greater than 0.05, indicating that the differences between the groups are not statistically meaningful. On the pretest, they actually performed the same.

Table 3

Descriptive Statistics (Language Achievement Post-test of Both Groups)

Groups	N	Means	Std. Deviations	Std. Error Means
Offline Group	31	14.03	1.83	.32
Online Group	32	16.25	1.98	.350

According to the descriptive data in the aforementioned table, the average on the language achievement post-test for the offline group was 14.03 whereas it was 16.25 for the online group.

Table 4

Inferential Statistics (Language Achievement Post-test of Both Groups)

		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
scores	Equal variances assumed	.37	.54	-4.60	61	.00	-2.21	.48
	Equal variances not assumed			-4.60	60.87	.00	-2.21	.48

Table 4 illustrates that there are meaningful differences between the offline and online participants at (p0.05). Actually, the online participants outperformed the offline participants on the post-test of language achievement. There is a substantial difference in the performances of the two classes on the language accomplishment post-tests in favor of the online group because the Sig value (.00) is less than.05.

Table 5

Paired Samples Test (Language Achievement Pre and Post-tests of Each Group)

		Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	Offline Group pre - Offline Group post	-1.77	2.74	.49	-3.60	30	.00
Pair 2	Online Group pre - Online Group post	-4.75	3.86	.68	-6.96	31	.00

Table 5 demonstrates that Sig (.00) is less than 0.05; consequently, there is a difference between the performances of the offline group before and after the intervention. Similarly, the discrepancies between the pre-test and post-test for the online class are notable as Sig (.00) is lower than 0.05.

Table 6

Descriptive Statistics (WTC Pre-test of Both Groups)

	Groups	N	Mean	Std. Deviation	Std. Error Mean
Scores	Offline Group	31	33.83	4.43	.79
	Online Group	32	34.90	4.61	.81

According to Table 6, the mean score for the offline class is 33.83, whereas the mean score for the online class is 34.90. On the WTC pre-test, it appears that both classes performed comparably well. To see if there were meaningful differences amongst the WTC posttests of the two classes, an independent samples t-test was done on the following table:

Table 7

Inferential Statistics (WTC Pre-test of Both Groups)

	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
scores	Equal variances assumed	.04	.82	-.93	61	.35	-1.06	1.14
	Equal variances not assumed			-.93	60.99	.35	-1.06	1.14

Table 7 shows that the Sig value is.35, which is more than 0.05, indicating that there were no variations in the WTC pre-test scores between the two groups. In actuality, they had the same performances before the treatment.

Table 8

Descriptive Statistics (WTC Post-test of Both Groups)

	Groups	N	Means	Std. Deviations	Std. Error Means
Scores	Offline Group	31	36.87	5.50	.98
	Online Group	32	42.56	9.53	1.68

Table 8 displays the descriptive outcomes from the WTC post-test for both classes. The means for the online and offline groups are 42.56 and 36.87, respectively. The online group appeared to perform better than the offline class on the WTC post-test.

Table 9

Inferential Statistics (WTC Post-test of Both Groups)

		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
scores	Equal variances assumed	16.36	.00	-2.88	61	.00	-5.69	1.96
	Equal variances not assumed			-2.91	49.90	.00	-5.69	1.95

The differences amongst the WTC post-tests of the online and offline classes are statistically substantial based on Table 9's Sig value, which is.00 and less than.05. Actually, the online group outstripped the offline class in the WTC post-tests. The benefits of the treatment (online evaluation) might be credited for this improvement.

Table 10

Paired Samples Test (WTC Pre and Post-tests of Each Group)

		Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	Offline Group pre - Offline Group post	-3.03	5.74	1.03	-2.94	30	.00
Pair 2	Online Group pre - Online Group post	-7.65	10.74	1.89	-4.03	31	.00

Based on the outcomes, the Sig value (.00) is lower than 0.05; so, there are substantial differences amongst the pre-test and post-test for the offline group. Similarly, there are significant variances amongst the pre-test and post-test for the online class.

Table 11
Descriptive Statistics (Test Anxiety Pre-test of Both Groups)

	Groups	N	Means	Std. Deviations	Std. Error Means
Scores	Offline Group	31	59.90	6.86	1.23
	Online Group	32	60.53	6.16	1.08

Table 12 shows that the average for the offline class is 59.90, while the average for the online class is 60.53. On the pre-test for test anxiety, it looks that the subjects in both groups received approximately identical results.

Table 12
Inferential Statistics (Test Anxiety Pre-test of Both Groups)

		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
scores	Equal variances assumed	.75	.38	-.38	61	.70	-.62	1.64
	Equal variances not assumed			-.38	59.84	.70	-.62	1.64

The inferential statistics for the two groups on the pre-tests for test anxiety are displayed in Table 12. The differences between the groups' pre-tests are not significant at (p0.05) since Sig (.70) is more than 0.05. In actuality, both groups conducted equally well on the test anxiety pre-tests.

Table 13:
Descriptive Statistics (Test Anxiety Post-test of Both Groups)

	Groups	N	Means	Std. Deviations	Std. Error Means
Scores	Offline Group	31	62.41	8.16	1.46
	Online Group	32	68.59	15.56	2.75

As depicted in the above table, the offline and online groups' respective mean scores are 62.41 and 68.59. On the post-tests for test anxiety, it appears that the online participants conducted better than the offline participants.

Table 14
Inferential Statistics (Test Anxiety Post-test of Both Groups)

	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
scores	Equal variances assumed	11.11	.00	-1.96	61	.00	-6.17	3.14
	Equal variances not assumed			-1.98	57.18	.00	-6.17	3.11

The inferential statistics of the two classes on the text anxiety post-tests are displayed in Table 14. The differences amongst the classes are substantial at (p0.05) as the Sig value (.00) is lower than 0.05. After the treatment, the online group did better than the offline group.

Table 15

Paired Samples Test (Test Anxiety Pre and Post-tests of Each Group)

		Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	Offline Group pre - Offline Group post	-2.51	5.98	1.07	-2.34	30	.02
Pair 2	Online Group pre - Online Group post	-8.06	13.89	2.45	-3.28	31	.00

Table 15 compares the outcomes of the pre- and post-tests for test anxiety for the two classes. According to the table, the offline group's pre- and post-test outcomes differ significantly. The pre- and post-test outcomes of the online group show a similar, substantial difference. According to the findings, on all three post-tests, both groups showed improvement; nevertheless, the online group outperformed the offline group. Here, it is implied that for EFL students learning English, online testing is preferable to offline testing.

Discussion and Conclusion

Based on the obtained results, the post-tests of the online group and the offline group showed a meaningful difference. On the three post-tests of the current research, the OA group actually outperformed the offline assessment group. The acquired results are consistent with those of Brown and Lee (2015), whose findings indicated that the OA condition had a favorable impact on students' WTC and enhanced participation and engagement in communicative activities.

Additionally, Chapelle and Voss (2016) found that OA produced a less tense and stressful testing atmosphere for language learners, which is consistent with our findings. Our findings were further supported by Wang and Chen's (2018) assertion that there were variations in the motivational and engagement impacts of online versus offline evaluation. Additionally, Brown and Lee (2015), who confirmed the impacts of online evaluation on EFL students' language achievement, support the findings of this survey. They claimed that the increased possibilities for practice and feedback offered by the OA had a beneficial impact on language achievement.

The findings of this examination also accord with those of Smith and Johnson (2017), who looked at how online and offline evaluation affected test-taking anxiety. Their findings suggested that OA, which offers students a more cozy and familiar atmosphere, could minimize test-taking anxiety. Additionally, Chen and Wang (2018), who confirmed the influences of OA on EFL students' WTC, concur with the results. Our outcomes concur with those of Johnson and Smith

(2016), who demonstrated the value of OA in terms of language proficiency, test anxiety, and communicative openness.

The results of this investigation are in keeping with Harasim's (2012) theory of online collaborative learning, which established the benefits of using CALL and online instruction in learning a language. According to the hypothesis, students may work together to resolve their problems, which can enhance their ability to learn English. Online collaborative learning employs the tools offered by the Internet to offer learning settings that promote teamwork and knowledge creation. Additionally, Siemens' (2005) connectivism hypothesis, which claims that learners know when utilize online teaching and make connections with their peers, supports our findings. Moreover, our outcomes are indorsed by the Social Constructivism theory explaining learning and teaching as intricate interactive social phenomena among students and teachers. Based on this theory, the teachers provide a social situation in which the learners can amass or build with others the knowledge essential to solve the problems. According to this assumption, learning is a series of practical social experiences in which students learn by performing, cooperating, and reflecting with others.

The better performance of the online group over the offline group can be attributed some advantages of online. Online tests can aid instructors in checking the correctness of the responses and delivering answers in real-time in less time (Bahari, 2021). OA tests are essential to the success of eLearning, and it is obvious that they have several advantages over offline mode. However, given that most education nowadays is self-paced and tailored to the individual needs of learners, evaluations likewise need to be dynamic, individualized, and perfect to fit their needs (Helfaya, 2019).

Online examinations also offer the simplicity and efficiency needed in today's changing economy. Without being restricted by time or place, the pupils can complete evaluation assessments at their own pace and on their own schedule. Additionally, teachers may simply keep track of student progress, create tests, and quickly mark quizzes and other assessments by using digital eLearning assessment tools. As a result, assessments are made to meet the requirements of the instructors and pupils. The flexibility and ease of online examinations are their main advantages. For the majority of students, the evaluation is perfect because it can be finished at any time and from any location utilizing digital devices. This makes education more individualized and takes into account different timetables and learning requirements (Noorbehbahani, 2022).

Online tests can provide students with immediate feedback to help them understand their strengths and shortcomings in real-time because the answers are instantaneously stored and automatically assessed. As a result, OAs have a number of advantages for students, including promoting self-reflection, prompt course adjustment, and a stronger grasp of the material (Marti-Ballester & Meo, 2020).

Digital assessment tools are frequently enhanced with interactive aspects to make them more interesting, such as multimedia information, gamification components, immersive experiences like interactive quizzes, timed tasks, and others. Online progress monitors for students and other real-time data, including time spent, grades, areas for growth, etc., are features of online examinations that teachers can access. These cutting-edge reports assist teachers in giving thorough direction and modifying their lesson plans to serve the requirements of each learners. To assist learners with specific needs, online tests can be made available with text-to-speech capabilities, scalable fonts, or alternate formats. Digital evaluation technologies may need an initial investment, but over time they can drastically lower assessment expenses.

OAs are more secure as they are conducted on secure platforms that prevent cheating and other forms of malpractice. OAs can also use biometric authentication to ensure that the correct person is taking the exam. The benefits mentioned for the OAs can be the reasonable justifications for the gained outcomes of the present investigation.

To conclude, this research examined the influences of OA and offline assessments on WTC, test taking anxiety, and language achievement among Iraqi intermediate EFL students. The gained outcomes indicate that using the OA was more effective than the offline assessment for EFL learners to develop their WTC and language achievement and to reduce their test taking anxiety.

The following are some helpful advice in relation to the research's findings. To assist EFL students in furthering their English learning, EFL teachers should incorporate online evaluation as well as other online tools and platforms into their lessons. Additionally, it is advised that EFL instructors and teachers increase their knowledge of anxiety-provoking situations, their impacts on students, and the methods to lessen them. This calls for the inclusion of test anxiety and related difficulties in both pre-service and in-service teacher education programs. Teachers ought to receive instruction on the causes of anxiety, how it affects learning, and coping mechanisms within the context of those courses.

More specially, instructors need to know how to support students who are struggling with issues like self-doubt, fear of failures, fear of rejection, bad experiences, and fear of criticism. Teachers should also teach their students how to create successful study habits and improve their test-taking abilities. They should also foster a conducive testing environment and regulate the procedures for administering the tests. On the other hand, educators ought to be aware that a reasonable amount of test anxiety is a cause to encourage students to work toward their goals, and that while anxiety is a predictor of learning motivation, performing well on tests is an agent that lessens test anxiety.

Teachers must be cognizant that test anxiety has both a cause and an effect related to fear of failing. Then, as EFL students struggle with their physical, emotional, and test-related issues, they want expert advice and therapy. Within this framework, school advisors and psychoanalysts should offer proficient services in the areas of eating and sleeping disorders, failure-related anxiety, time management, and test-related distractions, low motivation, and low self-esteem. Additionally, school counselors and guides should teach students how to unwind before exams. As a final point, given that parental expectations appear to be a source of worry, instruction plans ought to be developed to educate parents about the profound effect that their expectations have on their children.

The study includes a number of restrictions. First, 63 pupils who were enrolled in a private English school in Iraqi took part in the investigation. The scope of the research was restricted to the quantitative data gained from pre-tests and post-tests. The participants of the research were restricted to the male learners with the ages of 16 to 28. We included only urban students who had easy access to the technology; working on rural students might produce different results.

We offered some suggestions for upcoming investigations. First, since the results of the current study experimentally show the factors relating to WTC, language proficiency, and test anxiety, next studies can analyze each of those elements in correlational and descriptive investigations. Second, as the outcomes were gained by using quantitative tools; next research can apply qualitative tools including observations and interviews to get more valid results. Future studies can compare and contrast the influences of online and offline assessments on other English language sub-skills and skills. Additionally, experimental investigations are required to

contextualize the ways in which and why plausible elements may affect students' reactions to test anxiety as well as to comprehend the connections between test anxiety levels and, external, internal, and demographics factors.

Funding

This study is supported via funding from Prince Sattam Bin Abdulaziz University Project Number (PSAU 2023 /R/1444).

References

Abduh, M. Y. M., (2021). EFL teachers' online assessment practices during the COVID-19 pandemic: Changes and mediating factors. *Asian EFL Journal, 28*(1.1), 26–46. DOI:10.1007/s40299-021-00589-3

Alemu, B.M., & Feyssa, T. (2020). The relationship between test anxiety and academic achievement of grade ten students of Shirka Woreda, Oromia Regional State, Ethiopia. *African Educational Research Journal, 8*(3), 540-550. DOI: 10.30918/AERJ.83.20.097

Al-Samiri, R. A. (2021). English language teaching in Saudi Arabia in response to the COVID-19 pandemic: challenges and positive outcomes. *Arab World English Journal*, *(S1)*, 147-159. https://doi.org/10.24093/awej/covid.11

Amin, H. A. A., Khalil, H., Khaled, D., Mahdi, M., Fathelbab, M., & Gaber, D. A. (2021). Case item creation and video case presentation as summative assessment tools for distance learning in the pandemic era. *Medical Journal Armed Forces India*, *77*, 466–474. DOI: 10.1016/j.mjafi.2021.05.018 .

Alruwais, N., Wills, G., & Wald, M. (2018). Advantages and challenges of using assessment. *International Journal of Information and Education Technology*, *8*(1), 34–37. DOI:10.18178/ijiet.2018.8.1.1008

Alsadoon, H. (2017). Students' perceptions of e-assessment at Saudi electronic university. *Turkish Online Journal of Educational Technology-TOJET*, *16*(1), 147–153. ID: 67259787

Aydin, S. (2009). Test anxiety among foreign language learners: A review of literature. *Journal of Language and Linguistics Studies, 5*(1), 33-49. https://files.eric.ed.gov/fulltext/ED511092.pdf

Aydın, S., Yavuz, F., & Yeşilyurt, S. (2006). Test anxiety in foreign language learning. Balıkesir University. *Journal of Social Sciences Institute, 9*, 145- 160. https://dergipark.org.tr/tr/download/article-file/863782

Bahari, A. (2021). Computer-assisted language proficiency assessment tools and strategies. *Open Learning: The Journal of Open, Distance and e-Learning*, *36*(1), 61–87. https://doi.org/10.1080/02680513.2020.1726738

Brown, H. D., & Lee, H. (2015). The effects of online and offline assessment on willingness to communicate in a second language. *Language Teaching Research, 19*(3), 279-297. DOI:10.21608/mjat.2018.104629

Brown, H. D., & Lee, H. (2015). The effects of online and offline assessment on language achievement. *TESOL Quarterly, 49*(4), 733-756. Doi: 20.1001.1.24763187.2021.10.2.6.2

Care, E., Kim, H., Vista, A., & Anderson, K. (2018). *Education system alignment for 21st century skills: Focus on assessment.* Center for Universal Education at the Brookings Institution.

Chen, L., & Wang, Y. (2018). Willingness to communicate in online and offline assessment contexts. *System, 76,* 1-10. DOI:10.29140/jaltcall.v14n1.223

Chapelle, C. A., & Voss, E. (2016). The effects of online and offline assessment on test taking anxiety in language learners. *Language Testing, 33*(2), 169-191. Retrieved from https://ejournals.bc.edu/index.php/jtla/article/view/1648

Dinga J. N., Mwaura A. M., and Ng'ang'a M. W. (2018). Relationship between achievement goal orientation and academic achievement among form three students in Kiambu County, Kenya. *International Journal of Education and Research, 6*(4): 53-68. Doi:10.4103/0976-0105.141942

Fitriyah, I., & Jannah, M. (2021). Online assessment effect in EFL classroom: An investigation on students and teachers' perceptions. *Indonesian Journal of English Language Teaching and Applied Linguistics, 5*(2), 265-284. DOI:10.21093 /ijeltal.v5i2.709

Gürsoy, E., & Akin, F. (2013). Is younger really better? anxiety about learning a foreign language in Turkish children. *Social Behavior and Personality: An International Journal, 41,* 827-841. DOI: https://doi.org/10.2224/sbp.2013.41.5.827

Habibullah, S., & Ashraf, J. (2013). Factors affecting academic performance of primary school children. *Pakistan Journal of Medical Research, 52*(2), 76-92. Retrieved from https://www.scholarexpress.ne

Harasim, L. (2012). *Learning theory and online technologies.* New York: Routledge/Taylor and Francis.

Helfaya, A. (2019). Assessing the use of computer-based assessment-feedback in teaching digital accountants. *Accounting Education, Taylor and Francis, 28* (1), 69-99. https://doi.org/10.1080/09639284.2018.1501716

Jin, Y., & Zhang, L. J. (2019). A comparative study of two scales for foreign language classroom enjoyment. *Percept. Mot. Skills, 126*, 1024–1041. doi: 10.1177/0031512519864471

Johnson, R., & Smith, A. (2016). A comparative study of online and offline assessment on language achievement, test taking anxiety, and willingness to communicate. *Language Testing, 33*(2), 189-212. doi:10.4304/jltr.5.5.971-975.

Karaoglan-Yilmaz, F. G., Ustun, A. B., & Yilmaz, R. (2020). Investigation of pre-service teachers' opinions on advantages and disadvantages of online formative assessment: an example of online multiple-choice exam. *Journal of Teacher Education and Lifelong Learning*, *2*(1), 1–8. Retrieved from https://dergipark.org.tr/tr/pub/tel

Kavakci, O., Semiz, M., Kartal, A., Dikici, A., & Kugu, N. (2014). Test anxiety prevalence and related variables in the students who are going to take the university entrance examination. *Dusunen Adem the Journal of Psychiatry and Neurological Sciences, 27*(4), 301- 307. DOI:10.5350/DAJPN2014270403

Kocdar, S., Karadeniz, A., Peytcheva-Forsyth, R., & Stoeva, V. (2018). Cheating and plagiarism in e-assessment: Students' perspectives. *Open Praxis*, *10*(3), 221–235. DOI: https://doi.org/10.5944/openpraxis.10.3.873

Lee, J.S., & Lee, K. (2020). Affective factors, virtual intercultural experiences, and L2 willingness to communicate in in-class, out-of-class, and digital settings. *Language Teaching Research, 24*(6), 813-833. https://doi.org/10.1177/1362168819831408

Liu, Y., Zhang, M., Zhao, X., & Jia, F. (2021). Fostering EFL/ESL students' language achievement: The role of teachers' enthusiasm and classroom enjoyment. *Frontiers in Psychology, 12*, 1-6, doi: 10.3389/fpsyg.2021.781118

Luoma, S. (2017). The impact of online and offline assessment on language achievement: A meta-analysis. *Language Learning, 67*(2), 341-372. https://doi.org/10.3389/fpsyg.2023.12120567

MacIntyre, P. (2020). Expanding the theoretical base for the dynamics of willingness to communicate. *Studies in Second Language Learning and Teaching, 10*(1), 111-131, doi: 10.14746/ssllt.2020.10.1.6.

Malin, M., & Akimov, A. (2020). When old becomes new: a case study of oral examination as an online assessment tool. *Assessment and Evaluation in Higher Education*, *45*(8), 1205–1221. https://doi.org/10.1080/02602938.2020.1730301

Marti-Ballester, C. P., & Meo, F. (2020). Effects of the perceptions of online quizzes and electronic devices on student performance. *Australasian Journal of Educational Technology, 36*(1), 111–125. https:// www. resea rchga te. net/ publi cation/ 33396 7647

Mohamadi, M., Alishahi, Z., & Soleimani, N. (2014).A study on test anxiety and its relationship to test score and self-actualization of academic EFL students in Iran. *Procedia - Social and Behavioral Sciences, 98,* 1156 – 1164. https://doi.org/10.1016/j.sbspro.2014.03.529

Morgan, H. (2020). Best practices for implementing remote learning during a pandemic. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *93*(3), 135–141. https://doi.org/10.1080/00098655.2020.1751480

Munezane, Y. (2014). *A structural equation model and intervention study of individual differences, willingness to communicate and L2 use in an EFL classroom*. (Unpublished Doctoral Dissertation, Temple University, U.S.A.).

Noorbehbahani, F., Mohammadi, A., & Aminazadeh, M. (2022). A systematic review of research on cheating in online exams from 2010 to 2021. *Education and Information Technologies*. https:// doi. org/ 10. 1007/ s10639- 022- 10927-7

Novitasari, A., Wijayati, P. H., & Roekhan, R. (2020). Asynchronous digital-based learning transformation of speaking skill through instagram. *Journal DaFIna-Journal Deutsch Als Fremdsprache in Indonesien*, *4*(2), 33-48. DOI: http://dx.doi.org/10.17977/um079v4i22020p6-12

Nurhasanah, S., & Sobandi, A. (2016). Learning interest as determinant student learning outcomes. *J. Pend. Manaj. Perk. 1*, 128–135. doi: 10.17509/jpm. v1i1.3264

Oluoch, J. N., Aloka, J. O., & Odongo, B. C. (2018). Test anxiety beliefs as predictor of students' achievement in chemistry in public secondary schools in Kenya. *International Journal of Psychology and Behavioral Sciences, 8*(4), 70-76. ID: 189464359.

Or, C., & Chapman, E. (2022). Development and validation of an instrument to measure online assessment acceptance in higher education. *British Journal of Educational Technology, 53* (4), 977-997, doi: 10.1111/bjet.13180.

Richmond, G., Salazar, M. del C., & Jones, N. (2019). *Assessment and the future of teacher education.* SAGE Publications Sage CA: Los Angeles, CA.

Sarason, I.G. (1984). *The test anxiety scale: Concept and research.* Washington D.C.: Hemisphere Publishing Corp.

Shaw, A., Liu, O. L., Gu, L., Kardonova, E., Chirikov, I., Li, G., & Guo, F. (2020). Thinking critically about critical thinking: validating the Russian HEIghten® critical thinking assessment. *Studies in Higher Education*, *45*(9), 1933–1948. ID2-s2.0-85074051997

Shen, S., Mazgutova, D., & McCray, G. (2020). Exploring classroom willingness to communicate: the role of motivating future L2 selves. *International Journal of Educational Methodology, 6*(4), 729-743, oi: 10.12973/ijem.6.4.729.

Sheybani, M. (2019). The relationship between EFL learners' willingness to communicate (WTC) and their teacher immediacy attributes: a structural equation modelling. *Cogent Psychology, 6*(1), 1-14, doi: 10.1080/23311908.2019.1607051.

Sick, J. R., & Nagasaka, J. P. (2000). *A test of your willingness to communicate in English.* Unpublished questionnaire.

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning, 2*, 3-10. http://www.itdl.org/Journal/Jan_05/article01.htm

Smith, J., & Johnson, L. (2017). Exploring the impact of online and offline assessment on test taking anxiety. *Journal of Educational Psychology, 109*(2), 256-270. Retrieved from http://www.jtla.org.

Syed, H., & Kuzborska, I. (2018). Dynamics of factors underlying willingness to communicate in a second language. *The Language Learning Journal, 4,* 1-20. https://doi.org/10.1080/09571736.2018.1435709

Veena, T., & Mahlawat, S. (2020). Assessment in digital age: An overview of online assessment tools in education. *New Paradigm in Business & Education*, *6*, 247-260. DOI:10.24059/olj.v27i1.3398

Wang, Y., & Chen, N. S. (2018). Exploring the effects of online and offline assessment on language learners' motivation and engagement. *Computer Assisted Language Learning, 31*(1-2), 1-22. DOI:10.29333/ejmste/90992

Yashima, T. (2020). *L2 motivation and willingness to communicate*. The Palgrave Handbook of Motivation for Language Learning, pp. 203-223.

Yashima, T., MacIntyre, P.D., & Ikeda, M. (2018). Situated willingness to communicate in an L2: interplay of individual characteristics and context. *Language Teaching Research, 22*(1), 115-137. https://doi.org/10.1177/1362168816657851.

Yilmaz, F. G. K., Ustun, A. B., & Yilmaz, R. (2020). Investigation of pre-service teachers' opinions on advantages and disadvantages of online formative assessment: An example of online multiple-choice exam. *Journal of Teacher Education and Lifelong Learning, 2*(1), 10–19. https://dergipark.org.tr/en/pub/tell/issue/52517/718396

Yoestara, M., Putri, Z., Keumala, M., & Idami, Z. (2020). Pre-service English teachers' perception towards online assessment method. *International Journal of Education, Language and Religion, 2*(1), 1–10. DOI:10.35308/ijelr.v2i1.1933

Yüksel, H. S., & Gündüz, N. (2017). Formative and summative assessment in higher education: Opinions and practices of instructors. *European Journal of Education Studies*, *4,* 65-79. DOI: http://dx.doi.org/10.46827/ejes.v0i0.904

Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education, 16*(4), 323-342. https://doi.org/10.1207/S15324818AME1604_4