

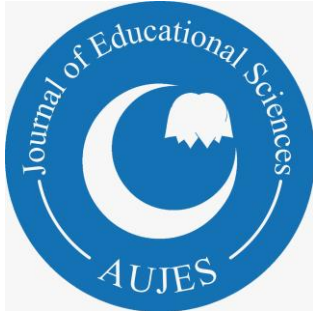
Article History

Received: 21.12.2023

Accepted: 25.03.2024

Available online: 30.06.2024

Article Type: Research Article




ADIYAMAN UNIVERSITY
Journal of Educational Sciences
(AUJES)

<https://dergipark.org.tr/tr/pub/adyuebd>

The Opportunities and Burdens of Online Test Processes

Akın Gürbüz¹, Bekir Canlı², Mehmet Abi³

¹ Vilnius University, Philology Faculty 

² Kahramanmaraş Sütçü İmam University, Department of Translation and Interpreting 

³ Muğla Sıtkı Koçman University, Sıtkı Koçman School of Foreign Languages 

To cite this article:

Gürbüz, A., Canlı, B. & Abi, M. (2024). The opportunities and burdens of online test processes. *Adiyaman University Journal of Educational Sciences*, 14(1), 1-9.

The Opportunities and Burdens of Online Test Processes

Akın Gürbüz^{1*}, Bekir Canlı², Mehmet Abi³

¹ Vilnius University, Philology Faculty

² Kahramanmaraş Sütçü İmam University, Department of Translation and Interpreting

³ Muğla Sıtkı Koçman University, Sıtkı Koçman School of Foreign Languages

Abstract

English language testing has been challenged by the COVID-19 pandemic as it required the test developers to instantly develop and implement online language tests in their unique contexts. In this regard, this study aims to investigate EFL language test writers' reflections on their planning, development, administration, reporting, and reflection processes of a test in the emergency remote teaching environment and display the potential opportunities and burdens these processes have brought about. The data were collected by utilizing semi-structured interviews with nine test writers who had testing experience in language preparatory classes at three state universities in Türkiye. The results indicated that the test writers were mainly challenged by ensuring the security of the exams, adapting question types to online platforms, and dealing with technical problems which resulted in additional workload. However, amidst these challenges, the study also identified notable opportunities, including the practical application of exam procedures, the digital transformation of exams with a variety of resources, item banking, ecological benefits, and fostering professional development. Overall, this compulsory experience during the pandemic indicated numerous conveniences to enhance language testing and contribute to the development of robust language assessment and testing in the future.

Key words: Language testing, language assessment, remote teaching, online testing, testing processes.

Introduction

The COVID-19 pandemic crisis caused a global transition to remote teaching and learning. The education domain globally, including in Türkiye, experienced a rapid shift from face-to-face to remote teaching, learning, and assessment (Güngör & Güngör, 2021) resulting in several issues due to insufficient preparation and infrastructure problems (Şenel & Şenel, 2021). Contrary to what would have been in an educational context, the transition process to remote teaching was unplanned and rushed, leaving many institutions to implement through trial and error (Naqv & Zehra, 2020).

The transition to emergency remote teaching affected the testing processes from several aspects. The long-term effects of this will undoubtedly leave a lasting impact on classroom-based assessment. Therefore, understanding teachers' online assessment practices and the factors that influence these practices is critical not only for assessment trainers but also for language educators and policymakers in applied linguistics (Zhang, Yan, & Wang, 2021). In this respect, understanding and adapting to the challenges and opportunities posed by online testing processes in teaching English as a foreign language (EFL hereafter) can pave the way for enhanced language learning experiences and evaluation methodologies.

The integration of online testing processes in EFL has brought forth a myriad of opportunities and challenges for educators and learners alike. As technology continues to transform the landscape of education, the use of online assessments has gained momentum, presenting educators with new possibilities to enhance learning experiences. However, amidst these opportunities lie unique challenges that demand careful consideration and adaptation to ensure the efficacy and fairness of the assessment process. In this article, we delve into the various opportunities and challenges posed by online testing in the field of teaching English as a foreign language. By examining the remote testing processes and potential benefits and obstacles from the perspectives of EFL test writers, we aim to shed light on ways to maximize the potential of online testing for effective language learning and evaluation. Consequently, there is a need for a methodological approach to understanding emergency remote online teaching and learning challenges and opportunities regarding test *planning, development, administration, reporting, and reflections*. In this regard, the primary focus of the present study is to investigate:

* Corresponding Author: Akın Gürbüz, akingurbuz@mu.edu.tr

- the way the EFL test writers reacted to remote teaching in terms of language test development,
- the similarities or differences between online and face-to-face EFL testing processes, if any,
- the opportunities brought about and/or challenges posed by online language test development.

Review of the Literature

Principles of Language Testing

To construct an effective language test, several language assessment principles and qualities need to be followed (Bachman & Palmar, 1996; Brown, 2004). One of the essential principles of language testing is known as validity, i.e., a test should measure what the test developers intend to measure. In other words, a valid test should evaluate the program's outcomes in line with the program's objectives. For example, a language test is supposed to examine test takers' language ability. Another critical principle of language testing is described as reliability. Stoyhoff and Chapelle (2005) discussed the reliability of a test "as the consistency reflected in test scores or as the absence of error in test scores." (p. 141). Reliability can also be defined as trust to inst results or consistency among characteristics of the testing situation (Bachman & Palmar, 1996). The third principle of language testing is accepted as authenticity. Bailey (1996) explained authentic test tasks are supposed to be related to real-life everyday situations. The fourth principle of language testing is acknowledged as practicality: "the adequacy of the available resources for the design, development, use, and evaluation of the test." (Stoyhoff & Chapelle, 2005, p. 144). The fifth principle of language testing is identified as washback which is elucidated as the positive or negative effect of the test on stakeholders. The sixth quality of language testing is defined as interactiveness: "the expected extent of involvement of test takers' knowledge and interest and their communicative language strategies in accomplishing a test task." (Stoyhoff & Chapelle, 2005, p. 143).

Test Development Processes

The test development process is regarded as a challenging task for test developers. Downing and Haladyna (2011) suggest a concise framework for test developers to facilitate test organization. The framework consists of 12 steps to develop a successful test. In this regard, Downing's systematic framework for effective test development may be useful for test developers to maximize test construction. Although all the steps are essential to a certain degree for creating effective tests, some of them may be skipped based on the test objectives. Applying these steps systematically as developing language tests may contribute to the validity, which is considered the essence of testing. Five themes emerged from this framework for the present study: planning, development, administration, reporting, and reviewing.

According to Downing and Haladyna (2011), the test planning phase commences with defining objectives, and in this pivotal stage tasks critical for a successful test are outlined. Remarkably, aligning test content with objectives ensures its success through content validity. Test specifications encompass format, item count, visuals, scoring rules, and time limits, while item selection involves choosing suitable question types based on the test's purpose. Moreover, maintaining validity in test item arrangement is essential, particularly considering the delivery mode (traditional or online). This phase concludes with careful printing and publishing, highlighting security and readability. Effective test administration necessitates standardized conditions, security protocols, and a watchful eye on cheating, particularly in online settings. Following test administration, scoring adheres to established measurement systems, requiring careful execution to preserve validity. Subsequently, essential result reports are generated for test-takers and stakeholders. A final review addresses any deficiencies, ensuring test readiness and providing recommendations for future implementations.

Opportunities Emerged from Online Testing

Although the emergent remote assessment due to the COVID-19 pandemic was a complicated and challenging issue in higher education at the same time, numerous opportunities have been reported in the related literature. The online assessment caused less anxiety among students, and the quality of the tests was considered good (Şenel & Şenel, 2021). In the same vein, Fitriyah and Jannah (2021) expressed that remote assessment strengthens flexibility regarding time and space, improves autonomous learning, cultivates preparation for online tests, and develops language assessment capacity. In addition, the online assessment was considered quite practical in delivering tests during the COVID-19 pandemic (Yulianto & Mujtahin, 2021). Moreover, the students found the remote assessment easy, enjoyable, and fun. Furthermore, thanks to remote assessment the students attained instant and direct feedback (Wibowo & Novitasari, 2021).

Kucherova and Ushakova's (2022) research demonstrates the successful application of various Moodle learning management systems online testing at the tertiary level. Students and instructors found online testing effective and relevant, with positive washback on education. The technology enables students to take the test anywhere, promoting flexibility and the comfort of their homes, and the content is not solely reliant on memory, providing a more comprehensive evaluation of students' language skills. Integrating various types of assessment

into General English courses contributed to self-driven learning and scaffolded students' learning throughout the term, while timely feedback facilitated continuous improvement. Specific technical settings, such as limited attempts and time restrictions, reduced cheating risks, making online testing a secure and valuable form of formal assessment.

Bui (2022) provides a comprehensive view of language testing and assessment in distance learning based on global and local literature. Online language testing offers numerous benefits, including promoting learner autonomy, evaluating student progress, and providing convenience. Likewise, Wibowo and Novitasari (2021) conducted a study on online assessment implementation and its impact on students' perceptions. Teachers utilized various tools such as Google Forms, Google Classroom, Quizzes, Edmodo, and Instagram for different skills like speaking, reading, and writing. The advantages of online assessment include auto-marking, immediate and quality feedback, reliable and valid measurement, efficiency, flexibility, and practicality.

Overall, the review of the literature highlights the benefits of online testing and its potential for future educational developments. Positive student perceptions highlight the benefits of well-prepared platforms, appropriate questions, clear instructions, and the flexibility of online assessments. The experiences with online testing may serve as a basis for future methodological developments, independent of lockdown restrictions, and may change the balance between online and offline formal assessment.

Challenges Posed by Online Testing

The COVID-19 pandemic has caused numerous difficulties for test developers in higher education. In their study, Guangul et al. (2020) emphasized that the COVID-19 pandemic not only influenced teaching and learning activities negatively but also affected assessment as it merged the existing problems of online assessment with the emergent COVID-19 pandemic in higher education institutions. The study reported academic misconduct, inadequate educational conditions, the amount of program output, and irresponsibility of learners in turning in homework as major challenges. In the same vein, Şenel and Şenel (2021) listed infrastructure problems and lack of online education experience for both instructors and learners as challenges the COVID-19 pandemic created for higher education institutions. Furthermore, the study revealed that ensuring fairness in test results, providing sufficient feedback, and covering the outcomes of the program was challenging.

The emergent online transformation in higher education due to the COVID-19 pandemic caused additional challenges concerning both instructors and learners in various settings. Montenegro-Rueda et al. (2021) listed instructors' shortage of education in online assessment for instructors and academic dishonesty for learners as major challenges. The study conducted by Hamad et al. (2021) also reveals significant shortcomings in online teaching across various domains. These include students' tendency to cheat during online examinations, the lack of class interactivity due to network problems, domestic distractions, and isolation, the inaccuracy of grades and results in reflecting students' true abilities, the demanding workload for instructors in planning and correcting, and difficulties in accurately assessing students' micro and macro skills as well as their motivation.

According to the research conducted by Meiantoni, Wiyaka, and Prastikawati (2021), the implementation of online assessment involves using online application-based media such as Google Classroom, Google Forms, WhatsApp groups, and Zoom Meetings for both formative and summative assessments in English classes. However, English teachers faced various challenges during the online assessment process, including unstable Internet connections for students, the need for additional time to prepare and adapt to the assessment media, varying student intelligence levels, lack of motivation and support, and limited understanding of the technology used for online assessments. Similarly, the findings of the study conducted by Kurniati and colleagues (2023) highlight comparable challenges faced by teachers at all levels while using online assessments, such as poor Internet connectivity, academic dishonesty, student discipline issues, and the lack of access to mobile phones.

In Wibowo and Novitasari (2021), teachers also faced challenges related to network and connection issues, designing assessments for specific skills, student control, and cheating. Nonetheless, students also expressed negative perceptions related to Internet access difficulties, distractions in noisy environments, and the time-consuming nature of online assessments. Moreover, Bui (2022) identifies several challenges, such as cheating, preference for selected response items, and concerns about validity and reliability, which increase teachers' workload. Recommendations for stakeholders, especially teachers and teacher education institutions, are presented to address these issues. Teachers are encouraged to utilize online formative assessment with prompt and personalized feedback, as well as opportunities for peer collaboration. These challenges significantly impacted the implementation of online assessments in English classes and underscored the need to address the challenges to improve the effectiveness of online teaching methods. Teacher education institutions should design training programs to enhance teachers' information and communications technology (ICT) competencies, emphasizing the effective use of technology in teaching and assessment. Additionally, flexible training methods and peer support groups are suggested to cater to teachers' diverse learning needs.

Method

Data collection

The research inquiry meticulously chose participants and the environment with careful consideration, aiming to optimize the chances of collecting extensive and valuable data from diverse viewpoints. The research encompassed the administration of semi-structured interviews with a group of nine EFL test writers who were affiliated with three state universities in Türkiye. The selection of these participant test writers was deliberate, as they were anticipated to provide valuable insights that may not have been accessible through alternative methods. As stated by Nunan (1992), the careful selection of participants based on their representativeness is essential when employing purposeful sampling techniques. Likewise, Patton (1990) characterizes purposeful sampling as a deliberate strategy that diverges from both probability sampling and convenience sampling, involving the intentional selection of specific settings, individuals, or events. Leavy (2014) emphasizes the benefits of purposeful sampling, noting that its nonprobability nature makes it particularly suitable for qualitative content analysis. This approach enables researchers to identify the most appropriate subjects for study without excluding individuals who are essential to the sample. Moreover, purposeful sampling enhances the chances of gathering in-depth, thorough, and significant data. It also provides the researcher with the flexibility to refine and adjust the sample as the study evolves, ensuring the collection of valuable information.

Participants

Participants for the study were chosen using a combination of purposive and convenience sampling methods. As the focus was on exploring the online testing practices of EFL test writers, an invitation email was expressly sent to the participants with whom one of the authors had a connection. Nine responded to the invitation and agreed to participate in the study. The participants' profiles, represented by pseudonyms, are provided in Table 1. Among the participants, six were female and three were male. While one of the participants held a doctorate, five had a master's degree and three had a bachelor's degree. The age range of the participants was between 30 and 55 years old, and their teaching experience ranged from 11 to 32 years. Their experience in testing units ranged from three to eight years.

Table 1. Profiles of the participants

Name	Gender	Age	Degree	Years of teaching	Experience in testing unit
Simon	Male	54	PhD	32	15
Kathrine	Female	52	BA	29	12
Ruth	Female	49	MA	27	6
Helen	Female	43	BA	20	8
Daphne	Female	42	MA	19	6
Celine	Female	40	BA	18	5
Henry	Male	32	MA	10	4
Nick	Male	30	MA	10	4
Julia	Female	33	MA	11	3

Data Analysis

To ensure a comprehensive and dependable comprehension of the data obtained from the semi-structured interviews, a meticulous qualitative content analysis was performed, delving into the details and nuances of the gathered information. As Weber (1990) explains, content analysis is a research methodology that employs a series of procedures to derive valid conclusions from textual data. In comparison to other research methods, content analysis presents several advantages. Firstly, it directly examines texts or transcripts, which are the outcomes of human communication and the cornerstone of social interaction. Additionally, this method is non-intrusive, preserving the integrity of the data and reducing the likelihood of participants reacting based on certain assumptions (Cohen et al., 2007; Stemler, 2001; Weber, 1990).

The coding of the raw data followed a cyclical process, which involved collaborating with a colleague and utilizing a cross-referencing approach to ensure interrater reliability at 81%. This resulted in a high level of agreement and accuracy in coding. Initially, the semi-structured interviews were recorded on video and transcribed verbatim, yielding a total of 34,790 transcribed words. Codes were then agreed upon by the researcher and the interrater, based on the dimensions of the study. Once the codes were established, both the researcher and the interrater independently organized the raw data from the interviews into themes and subthemes. At each stage of the coding process, the categorization of excerpts under relevant themes and subthemes was reviewed, revised, and ultimately finalized by referring back to the raw data. This rigorous process ensured a comprehensive understanding of the data and promoted consistency in the analysis.

Ethics Approval

The ethics application for the study was made on 13/10/2023, and the research was carried out with the approval of the Muğla Sıtkı Koçman University Ethics Commission dated 09/12/2023 and numbered 118.

Findings

The qualitative content analysis of the interviews revealed five themes regarding online testing processes namely test planning, test development, test administration, reporting test results, and reviewing test results. The frequency of subthemes that emerged from these themes is presented in Table 2. The semi-structured interviews and video-stimulated interviews were designed to elicit the EFL test writers' perceptions of the challenges and effectiveness of online testing and assessment procedures.

Table 2. The Frequency of Themes Emerged from Semi-structured Interviews

Opportunities	<i>f</i>	Challenges	<i>f</i>
practical and easy administration of tests	22	limited types of questions/tasks	16
online & softcopy resources	14	security issues	14
instant feedback	6	lack of feedback among test writers	10
convenient	4	technical problems	8
item banking & digital transformation of exams	4	lack of collaboration among test writers	2
statistical analysis of results	4	extra workload	2
ensuring learner privacy	4	unequal opportunities for learners	2
professional development in using technology	4	lack of assessment in speaking	2
online communication platforms	4		
ecologic benefits	2		
standardizing exam duration	2		
increasing objectivity	2		
enjoyable	2		
Total	74	Total	56

The participant test writers most frequently cited practical and easy administration of tests (22 times) as an opportunity that emerged from online testing. They found online testing more practical, especially in the planning and test administration phases as it required less workload and fewer physical requirements. Online administration of tests enabled test writers to avoid the workload of planning and preparing the physical conditions of tests. They stated that online administration of tests took the load of preparing the conditions of classrooms for exams and placing students into classrooms, printing and sorting exam papers, preparing invigilation lists and assigning instructors, dealing with planned or unplanned circumstances during the physical administration of tests, collecting exam papers at the end of the exams, and tasks alike which require physical exertion. Participant test writers also found online tests flexible and timesaving as they could be administered asynchronously, and even at the weekends without taking out from teaching time. For instance, one of the participants asserted that:

Julia: *Conducting an exam [face-to-face] used to take one whole school day, taking 5 or 6 teaching hours. Doing this at the weekend helped save time for the flow of the syllabus.*

Another significant subtheme that emerged from the data was related to the variety of online and soft copy resources and materials used in exam content (cited 14 times). The participants emphasized that they could benefit from more online and softcopy materials rather than mainly relying on hardcopy resources as they previously did. This has also led to the digital transformation of exams and item banking (cited 4 times) in institutions. Online testing and assessment allowed test writers to diversify exam content, while also necessitating limitations on the types of tasks and questions used in exams. This was noted in the interviews 12 times. While the test writers were able to use more digital materials such as pictures, videos, MP3s, and so on in their exam content, they also felt necessary to avoid those question types that might cause vagueness and ambiguity in the proper answer which might not match the answer set into the system. They claimed that manual checks of the exam papers allowed instructors to be able to act more on their initiatives during face-to-face administration of exams; however, using an online platform for the exams took this initiative and made it necessary to avoid using open-ended and/or re-write question types and use fixed-answer question types such as matching, true-false, multiple choice and so on. The participants stated that: All tables and figures should be embedded in appropriate areas within the document and centered. They should not exceed the page margins.

Daphne: *We do not have the flexibility in responses. We need to form fixed-answer questions like multiple-choice questions.*

Celine: *In the face-to-face exam, we could make use of all question types such as true-false, gap-filling, sentence completion, and so on. But online testing limited the range of question types with multiple-choice questions because of the applicability. This is a significant drawback for us [test writers].*

Some further benefits of online test administration were found to be instant feedback for students (cited 6 times), statistical analysis of results (cited 4 times), and ensuring learner privacy (cited 4 times). With the use of digital tools, learners can get immediate feedback and results upon the completion of a test. These digital platforms also provide the statistical analysis of the learner results, which provides the test writers with practical feedback about the exam they have prepared and administered, and microanalysis of the exam by conducting item analysis. Test writers could easily obtain data about the difficulty of the exams, how suitably it was leveled, and implications for further exams and/or question types. It is evident that the use of online testing platforms provided educators with resourceful data. To illustrate, one of the participants uttered:

Henry: *Online testing made it possible to obtain exam results automatically and instantly. We can also obtain statistical analysis of the results immediately after the exam.*

Regarding the opportunities that arose through online exam administration, the participant test writers further asserted that they found it more convenient (4 times) as they could work from home; they found it as an opportunity for professional development to use technology (cited 4 times) more efficiently; they had more opportunities to use online communication tools (cited 4 times) to inform the instructors about exams and receive feedback from them; and they found it more enjoyable (cited 2 twice) in practice; and online testing increased the objectivity of assessments due to the preference of fixed-answer questions, created a standardization in exam durations (cited 2 twice), and provided ecologic benefits (cited 2 twice) as they could save from paper consumption. To illustrate, the participants highlighted that:

Ruth: *I have realized that writing can be measured with different techniques. For example, I prepared quizzes and exams for my students as if I were playing a computer game at home. In this respect, it increased my creativity in crafting exams and contributed to my professional development.*

Simon: *Let me tell you the biggest advantage. Paper savings. We have saved toner for printing. These are financial savings. There was so much consumption that I think this was both economically and ecologically saving.*

In addition to limiting the question types, the participant test writers found online test administration challenging as they had concerns regarding security issues (14 times). No matter whether test takers are obliged to open cameras during exam administration (which was not the case for most state universities in the country), test writers felt that there was a huge gap in the security issues and that learners were taking unfair exam scores because they could cheat during the exams in any way. The test writers found this unjust and discouraging in exchange for the effort they put into the preparation of the exams. This indicates that test writers have significant concerns about the security measures taken for the administration of exams and feel unenthusiastic as a result. To illustrate, one of the participants asserted that:

Helen: *I am not complaining about the workload, however, is it worth all this effort? In other words, when there is no security in the exam, or rather, when we are not sure about it, not only me but all our instructors and even the administration. When we are not sure about this. We try so hard, but what if students cheat or pass undeservedly? The question mark is always in our minds. I am a little upset about this.*

Another noteworthy challenge that emerged from the analysis of the data concerned the lack of feedback among test writers (cited 10 times). Test writers complained that lockdowns reduced the amount of discussion they made over the exams they were preparing, and they could merely provide distant feedback via email, and/or audio/video calls. It seems evident that the idea of distant connections with colleagues even discouraged them from debating on the exams and reduced collaboration (cited twice). The participant test writers further claimed to be concerned about technical problems (cited 8 times) that occurred during the administration of exams. These challenges mainly stemmed from the lack of efficient infrastructure, learners' Internet connection problems, learners' lack of proper devices to accomplish the exam, and so on. By the same token, institutions received complaint letters from the learners; hence, test writers were supposed to administer make-up exams for those who were unable to complete the exam due to such technical problems. As a negative consequence of online testing, test writers had to prepare additional make-up exams for unpredicted technical problems. For example, one of the participants uttered:

Nick: Now, students themselves may experience technical problems. They may not have Internet connection, may have problems with their computer, or their connection may be lost during the exam, and so on. Then the students are naturally given the right to take a make-up exam.

Regarding the challenges posed by online test administration, the participants found it challenging since it brought extra workload (cited twice), created unequal distribution of opportunities (cited twice) among learners, and limited the assessment (cited twice), especially in speaking exams in terms of the evaluation of body language, mimes, gestures, and so on. They stated that online administration of exams brought about an additional step in the planning and preparation phases of exam development as it required uploading the exam content onto an online platform. This step required further caution and effort as some asserted that they had to reconstruct an examination because it was mistakenly shared with students before the due date. These were reported to be unpredicted aspects of online test administration. In a similar vein, test writers claimed that online administration of exams caused unjust circumstances among learners since they came from various socio-economic backgrounds. They believed that Internet access and devices being used were not equal among learners, which was creating disadvantaged groups.

Discussion and Conclusions

The focus of the study was to investigate how EFL test writers reacted to online testing, the potential differences in testing procedures compared to traditional testing, and the opportunities and challenges brought about by online testing. Regarding testing processes, the study revealed that online testing procedures did not significantly differ from traditional testing, with the main distinctions being in test administration and reporting results. The study highlighted that online testing offered greater convenience in test administration and result announcement. However, it also introduced additional burdens, such as the need to load and proof exam questions on online platforms. The findings further indicated that EFL test writers initially faced anxiety and uncertainty while adapting to online testing due to their lack of experience in this area. Nevertheless, they were able to overcome these apprehensions by gaining experience and exploiting incidents of professional development during the period. Despite the challenges posed by online testing, the research demonstrated that test writers were able to ultimately benefit from this demanding situation.

Surprisingly, unlike previous research on remote teaching and learning, this study revealed that the opportunities brought about by online test administration outweighed the challenges posed (Bui, 2022; Guangul et al., 2020; Hamad, et al., 2021; Kurniati et al., 2023; Meiantoni, et al., 2021; Şenel & Şenel 2021; Wibowo & Novitasari, 2021). The results indicated that EFL test writers found online testing convenient, practical, and easy to manage, which aligns with existing literature (Bui, 2022; Wibowo & Novitasari, 2021; Yulianto & Mujtahin, 2021). Similar to the emphasis on providing direct, immediate, and quality feedback in Wibowo and Novitasari (2021), and timely feedback for continuous improvement in Kucherova and Ushakova's (2022) work, online testing was found to offer instructors and learners immediate feedback and statistical data on learner performance by ensuring learner privacy, leading to positive outcomes of online testing.

A notable and distinctive finding of this study, setting it apart from previous research, revolves around the significance of the digital transformation of exams and the effective utilization of software resources. The results indicated that test writers placed significant emphasis on the use of online resources, banking the exams on digital platforms, and minimizing the use of paper and printing use, which were depicted as ecological benefits. Moreover, the study unveiled that test writers perceived online testing as a means to enhance the objectivity of tests, standardize the exam duration, enhance their technological skills, and make the testing process more enjoyable to conduct. Notably, these findings were consistent with earlier research (Şenel & Şenel, 2021; Wibowo & Novitasari, 2021).

Considering the challenges posed by online testing, the study indicated that the range and types of questions used were restricted by the online platforms employed, as suggested in Bui (2022) and Wibowo and Novitasari (2021). In contrast to the flexibility in time and place offered by online testing, it compelled the test writers to prioritize closed-ended question types and adjust the question formats they had previously used. Another major challenge posed by online testing revolves around security issues, which highlighted the EFL instructors' concerns regarding learners' tendencies to cheat and engage in academic dishonesty as previously noted by numerous scholars (Bui, 2022; Fitriyah and Jannah, 2021; Şenel & Şenel, 2021; Wibowo & Novitasari, 2021; Yulianto & Mujtahin, 2021). The results underpinned serious concerns about the accuracy of the scores obtained by learners in reflecting their actual abilities, which is leading test writers to believe that their efforts might be futile. Consistent with these findings, Yulianto and Mujtahin (2021) emphasized that online testing leads to an inaccurate assessment of learners' micro and macro skills, as well as their motivation.

Consistent with prior research (Meiantoni et al., 2021; Şenel & Şenel, 2021; Wibowo & Novitasari, 2021), another common challenge brought about by online test administration involved technical and network

problems resulting from infrastructure issues, which causes unequal opportunities for learners as the test writers highlighted in the study. In addition to the demanding workload for instructors in planning, test writers were compelled to prepare additional make-up exams for learners who lost connection or experienced technical problems during the test administration. As stated by Meiantoni et al. (2021), remote testing demanded extra time for test writers to prepare and adapt to the assessment media.

Another burden revealed in the study pertains to test writers' adaptation to online testing, although they were able to transform this challenge into an opportunity for professional development. As asserted by Meiantoni et al. (2021), the EFL test writers had a limited understanding of the technology used for online assessments due to the lack of experience in remote testing, as also observed by Montenegro-Rueda et al. (2021). A final remark of the study is that despite facilitating immediate feedback on learner performance, online test administration has led to a decrease in feedback and collaboration among test writers.

In summary, the study emphasizes that online testing, despite its challenges, offers numerous opportunities for educational institutions. The experience of the COVID-19 pandemic has underscored the need for these institutions to adapt to remote or flipped testing processes likely to be carried out in the future. To successfully achieve this, providing essential professional training for test writers and instructors, improving infrastructure, and prioritizing exam security measures are crucial steps to overcome potential challenges. By embracing these changes, educational institutions can harness the benefits and potential of online testing while effectively addressing its associated obstacles.

Recommendations

Considering the aforementioned conclusions, there are several areas for further exploration and research in the realm of online testing and its implications for educational institutions. Delving deeper into the specific strategies and methodologies for delivering professional training to test writers and instructors could yield valuable insights into enhancing the quality and reliability of online assessments. Additionally, investigating the technological aspects of infrastructure improvement, such as the development of user-friendly and secure online testing platforms, could contribute to the seamless implementation of remote or flipped testing processes. Exploring innovative approaches to exam security, including advanced authentication methods and anti-cheating measures, could further bolster the credibility and integrity of online assessments.

Furthermore, a comparative analysis of different modes of online testing, such as synchronous versus asynchronous formats, could shed light on their respective advantages and limitations. Long-term studies tracking the efficacy of educational institutions' adaptations to remote testing in the post-pandemic era could provide valuable data on the sustained benefits and challenges of these approaches. Lastly, examining the perspectives and experiences of both educators and students in these evolving contexts could offer a comprehensive understanding of the multifaceted dynamics involved in the transition to online testing.

Acknowledgments or Notes

We would like to express our heartfelt gratitude to Prof. Dr. Emrah Cinkara, whose invaluable insights and guidance have played a pivotal role in inspiring and shaping the research design for the study. His expertise and dedication have been instrumental in guiding our exploration of this important topic. We are deeply appreciative of his support and mentorship throughout this research endeavor.

References

- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1). Oxford University Press.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257-279. <https://doi.org/10.1177/0265532296013003>
- Brown, H. D. (2004). *Language assessment principles and classroom practice*. Oxford: Oxford University Press.
- Bui, T. H. (2022). A Review of Language Testing and Assessment in Online Teaching. *International Journal of English Linguistics*, 12(4), 54–65. <https://doi.org/10.5539/ijel.v12n4p54>
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). London and New York, NY: Routledge Falmer.
- Downing, S. M. & Haladyna, T.M., (Eds.). (2011). *Handbook of Test Development* (1st ed.). Routledge.
- 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. <https://doi.org/10.21093/ijeltal.v5i2.709>
- Guangul, F. M., Suhail, A. H., Khalit, M. I., & Khidhir, B. A. (2020). Challenges of remote assessment in higher education in the context of COVID-19: a case study of Middle East College. *Educational*

- Assessment Evaluation and Accountability*, 32(4), 519–535. <https://doi.org/10.1007/s11092-020-09340-w>
- Güngör, M. N., & Güngör, M. A. (2021). Reconceptualizing Testing in Times of COVID-19 from an Activity Theory Lens: A Narrative Inquiry of a Test-designer's Experiences. *The Asia-Pacific Education Researcher*, 30(6), 531–539. <https://doi.org/10.1007/s40299-021-00627-0>
- Hamad, M. M., Dafaallah, A. S., & Alhaj, A. A. (2021). The challenges of online English language teaching from EFL instructors' point of view in KKU, Tehama Campus during the COVID-19 period. *Journal of Language Teaching and Research*, 12(4), 543–556. <http://dx.doi.org/10.17507/jltr.1204.04>
- Kucherova, O., & Ushakova, I. (2022). Effectiveness of online testing in general English university courses from teacher and student perspectives. *Information Technologies and Learning Tools*, 87(1), 185–198. <https://doi.org/10.33407/itlt.v87i1.4812>
- Kurniati, D., Riyono, A., & Utomo, B. (2023). Challenges in online assessment: The case of Indonesian teachers. *International Journal of Social Science*, 2(5), 2063–2068. <https://doi.org/10.53625/ijss.v2i5.4811>
- Leavy, P. (Ed.). (2014). *Oxford library of psychology. The Oxford handbook of qualitative research*. Oxford University Press.
- Meiantoni, D., Wiyaka, W., & Prastikawati, E. F. (2021). Online Assessment in English Classroom: EFL Teachers' Practices and Challenges. *Journal of English Education and Linguistics*, 2(2), 26–36. <https://doi.org/10.56874/jeel.v2i2.532>
- Montenegro-Rueda, M., Luque-de la Rosa, A., Sarasola Sánchez-Serrano, J. L., & Fernández-Cerero, J. (2021). Assessment in higher education during the COVID-19 pandemic: A systematic review. *Sustainability*, 13(19), 10509. <https://doi.org/10.3390/su131910509>
- Naqvi, S., & Zehra, I. (2020). Online EFL emergency remote teaching during COVID-19, Challenges and innovative practices: A case of Oman. In *Arab World English Journal (AWEJ) Proceedings of 2nd MEC TESOL Conference*.
- Nunan D. (1992). *Research methods in language learning*. Cambridge University Press.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Sage Publications, Inc.
- Stemler, Steven. (2001). *An overview of content analysis. Practical assessment, research, and evaluation*. 7. [available online at] <http://pareonline.net/getvn.asp?v=7&n=17>
- Stoynoff, S., & Chapelle, C. A., (2005). *ESOL tests and testing: A resource for teachers and program administrators*. Alexandria, VA: TESOL.
- Şenel, S., & Şenel, H. C. (2021). Remote assessment in higher education during the COVID-19 pandemic. *International Journal of Assessment Tools in Education*, 181–199. <https://doi.org/10.21449/ijate.820140>
- Weber, R. P. (1990). Sage University paper series on quantitative applications in social sciences, No. 07-049. *Basic Content Analysis* (2nd ed.). Sage Publications, Inc.
- Wibowo, F. E., & Novita Sari, U. (2021). An Analysis of Online Assessment in Teaching English. *PROJECT (Professional Journal of English Education)*, 4(3), 521–529. <https://doi.org/10.22460/project.v4i3.p521-529>
- Yulianto, D., & Mujtahid, N. M. (2021). Online assessment during COVID-19 pandemic: EFL teachers' perspectives and their practices. *JET (Journal of English Teaching)*, 7(2), 229–242. <https://doi.org/10.33541/jet.v7i2.2770>
- Zhang, C., Yan, X., & Wang, J. (2021). EFL teachers' online assessment practices during the COVID-19 pandemic: Changes and mediating factors. *The Asia-Pacific Education Researcher*, 30, 499-507. <https://doi.org/10.1007/s40299-021-00589-3>