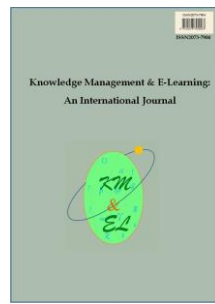

A systematic review of EFL online assessment in higher education: Effectiveness, attitudes, and challenges

Amr Selim Wannas

Aswan Heart Centre (AHC), Aswan, Egypt

Marwa M. AbdelMohsen

Arab Academy for Science, Technology, & Maritime Transport (AASTMT), Alexandria, Egypt



Knowledge Management & E-Learning: An International Journal (KM&EL)
ISSN 2073-7904

Recommended citation:

Wannas, A. S., & AbdelMohsen, M. M. (2025). A systematic review of EFL online assessment in higher education: Effectiveness, attitudes, and challenges. *Knowledge Management & E-Learning*, 17(3), 435–453.
<https://doi.org/10.34105/j.kmel.2025.17.020>

A systematic review of EFL online assessment in higher education: Effectiveness, attitudes, and challenges

Amr Selim Wannas* 

Magdi Yacoub Heart Foundation (MYF)
Aswan Heart Centre (AHC), Aswan, Egypt
E-mail: amrsilem8842@outlook.com

Marwa M. AbdelMohsen 

College of Language and Communication (CLC)
Arab Academy for Science, Technology, & Maritime Transport
(AASTMT), Alexandria, Egypt
E-mail: marwa.mohsen@aast.edu

*Corresponding author

Abstract: The global pandemic caused by COVID-19 has acted as a significant driver for the adoption of online educational practices, including instruction and assessment. Given the urgent need for an alternative to traditional assessment methods during the pandemic, various studies have explored this topic across different contexts to evaluate the impact, advantages, challenges, and learners' perspectives on online assessment. The current paper presents a systematic review that focuses on online assessment in teaching and learning English as a foreign language (EFL) in higher education. The review summarizes existing findings on the effectiveness of online assessment, student attitudes towards its use, and the challenges of online assessment. The PRISMA model was used for article screening, inclusion, and exclusion. Out of the screened articles, 19 were included in the review process. The findings indicate that online language assessment tools provide significant gains in developing various language skills and offer benefits in terms of speed, adaptability, interactivity, and learner autonomy. Nonetheless, there are notable challenges related to technology, pedagogy, and psychology that need to be addressed. To leverage the advantages of online language assessment, the review suggests approaches that can be used to address the most prominent challenges.

Keywords: Online language assessment; EFL education; Systematic review

Biographical notes: Amr Selim Wannas is a dedicated EFL/ESP instructor with over 12 years of experience in TESOL. He holds an MSc in English for Specific Purposes (ESP) and is a Cambridge CELTA-certified instructor. Amr has taught diverse age groups across various academic institutions, including Amideast Egypt and the Magdi Yacoub Heart Foundation. His exposure to different teaching practices has honed his teaching skills and broadened his perspective, enabling him to conduct research in language education. His research interests include TESOL and ESP.

Dr. Marwa AbdelMohsen is currently the Head of the Language and Translation Department at the College of Language and Communication at the AASTMT. With a rich academic background of over 25 years in higher education

institutions, she currently teaches Linguistics, Phonetics, Translation, and Language Assessment at both graduate and postgraduate levels. Dr. Marwa's research interests include language assessment, AI in language pedagogy and assessment, and educational technology.

1. Introduction

Technology has revolutionized the landscape of language education, especially during the global pandemic crisis. This significant change was driven by the rise of a new generation of computers and smartphones. Yet, the constant utilization of computers, mobile phones, and social media has ignited the interest of researchers and practitioners to develop the most effective techniques and applications in language teaching. This has been supported by various robust studies on the impact of technology on language acquisition (e.g., Chun et al., 2016; Shadiev & Yang, 2020). More recent trends in language teaching and learning include incorporating online platforms and eBooks, along with adopting various online assessment techniques.

2. Context and background

Language assessment is “*a systematic procedure for eliciting test and non-test data (e.g., a teacher checklist of student performance) for the purpose of making inferences or claims about certain language-related characteristics of an individual.*” (Purpura, 2016). It is carried out in a variety of educational and professional settings (e.g., large-scale international tests, curriculum-based progress tests and school-leaving examinations, certification assessments for employment purposes) and is frequently thought of as a related but independent disciplinary area from general educational assessment (Alderson et al., 2017).

2.1. Language assessment

There are four primary types of assessment employed in language education: diagnostic, formative, summative and integrative. Diagnostic assessment is usually used to gauge learners' existing knowledge and readiness prior to a given course (Crisp, 2011). Formative assessment encompasses tasks implemented throughout a course in conjunction with timely and appropriate feedback. The goal of these formative tasks is to promote the development of their performance in subsequent summative evaluations through feedback. The third type of assessment, summative assessment, is used to measure the overall learning progress and to fulfil certification requirements. The last type of assessment, integrative assessment, involves comprehensive assessments that require students to establish connections between knowledge and learning across different modules and/or subjects.

2.2. Online language assessment in higher education

Online assessment, also referred to as electronic assessment, takes place in various higher education institutions globally. Online assessment utilizes electronic devices to aid in constructing, presenting, storing, or analyzing student evaluation materials, responses,

scores, or feedback (Crisp, 2011). This mode of assessment is swiftly supplanting conventional examinations in higher education institutions and universities.

Due to the relative novelty of this domain for both educators and learners, online assessments may exhibit a higher frequency of measurement errors compared to traditional methods (Sievertsen & Burgess, 2020). Fitriyah and Jannah (2021) classify online assessment into various forms, including items for submission, automated assessment, online discussion, and web publishing. As shown in Table 1, these classifications encompass various subsections that outline the potential applications of online assessment (Fitriyah & Jannah, 2021).

Table 1
Forms of online assessment

Forms	Examples
1. Submission of items for assessment	• Essays • Reports • Reviews (Critical, Analytical) • Media (Image, Audio, Video, Presentation) • Multiple choice • Short answer
2. Automated assessment Quizzes Multimedia	• Calculation • Matching • Fill blanks • True/False • Drag & Drop • Simulations
3. Online discussion	• Forum (Case Analysis, Project Development) • Debates • Allocated roles (Lead, Summarize & Provoke) • Role plays
4. Web publishing	• E-portfolios • Webpages (Blogs, Wikis) • Shared documents (Google Docs)

Various studies reported mixed findings on the effectiveness of these online assessment forms in measuring the ILOs of language courses in higher education. Thus, a comprehensive view of the findings reported by studies in several language education contexts needs to be established. The current review attempts to fulfil this need by providing a thorough review of the studies that addressed this issue over the last decade.

2.3. Online language assessment tools

With the increased need for online education, various online language assessment tools that include systems, applications, or platforms have been developed to be used by educational institutions. Some of these tools focus on scoring, such as Automated Essay Scoring systems (AES) with or without feedback reports, and other tools provide a platform that teachers can use to generate and score language tests, or an application that can generate quizzes for formative assessment purposes. Among these tools are C-DA, Plickers, M-assessment, Quizziz, Socrative Quizzes application, and Moodle.

Computerized Dynamic Assessment (C-DA) serves as a diagnostic instrument that addresses fully internalized or in-progress second or foreign language skills (Kamrood et al., 2019). In generating scores, it capitalizes on Vygotsky's theoretical framework, leading directly to the design of customized instructional courses for the various types of learners (Ebadi & Saeedian, 2016). Mobile applications exhibited further personalized learning experiences in recent years, enabling learners to enhance their language skills through instant performance feedback. M-assessment, on the other hand, facilitated the knowledge evaluation process through making it more comfortable and leading to improved achievement in courses (Titova, 2019). In support of language assessment, Moodle platform plays a pivotal role in generating interactive randomized tests for formative and summative assessments, making it an essential tool for blended learning as well (Acar & Kayaoğlu, 2020; Schweighofer et al., 2019).

Several studies reported that AES systems have demonstrated tremendous ability to score learners' essays through utilizing Natural Language Processing (NLP), minimizing scoring time and cost, and increasing rater reliability (Bai et al., 2022; Hussein et al., 2019; Wong & Bong, 2019). Another renowned tool utilized frequently for creating quizzes is Google Forms. Google Forms are distinguished by their accessibility, efficacy, and efficiency (Maret'atush S et al., 2021; Radhaswati & Santosa, 2022). Similarly, Socrative Quizzes, which is a web-based platform, is known for its responsiveness, personalization and efficacy in formative EFL assessment (Alharbi & Meccawy, 2020; Kolisnyk et al., 2022). Other online applications, such as Kahoot!, operate learning on the concept of gamification through optimizing learners' motivation to take vocabulary and grammar tests, allowing for a break from traditional language classes (Hadijah et al., 2020; Wang & Tahir, 2020).

2.4. Attitude towards online assessment

Despite the reported practicality of online assessment, especially with the emergence of new factors that expand the need for distance education and assessment, the acceptance of this technology by learners represents an essential factor in the effectiveness of its use. Örnfelt (2023) clarified that the use of new educational technologies without monitoring the learners' attitude towards them can lead to a major knowledge gap.

In this regard, several scholars suggested models that can help in exploring the attitude of students and their perceived effectiveness of online assessment tools. In his systematic review of educational technologies, Granić (2023) highlights that the Technology Acceptance Model (TAM) is considered one of the most influential in investigating the attitude of stakeholders, including students and teachers, towards various emerging educational technologies. Over the years, TAM has undergone several modifications to incorporate various constructs related to the attitude towards using new technologies, such as perceived enjoyment, conformity behavior and self-esteem.

However, various studies reported mixed results in relation to the attitude of students towards the use of emerging assessment technologies, which suggests that there is a dire need for a comprehensive view of the students' perspective to ensure optimal application of the new technologies. The current systematic review attempts to fulfil this need by examining the results reported by several research studies that addressed this issue.

2.5. Research questions

While several review articles have explored online assessment in education, examining its prospects and challenges, such as Jurāne-Brēmane (2023), none have specifically delved into the realm of online language assessment in higher education. The current study aims to address this gap in the literature. Given the scarcity of research on the technological applications used in language assessment, the principal aim of this systematic review is to examine the effectiveness of online assessment, unveil learners' perceptions and attitudes towards this mode of assessment, and expound the challenges it faces as previously highlighted in the selected articles. The review addresses the following questions.

RQ1: How effective are online assessment tools/applications in evaluating and/or supporting EFL learners?

RQ2: What is the attitude of EFL learners towards online assessment?

RQ3: What are the challenges that face EFL online assessment?

3. Method

This review article employed the qualitative approach in applied linguistics research. It followed a process comprising seven key stages: identifying the review topic, formulating the research questions, establishing a protocol for article selections, conducting the search for the articles, selecting the articles based on the protocol, extracting the data and synthesizing the results, and discussing the outcomes of the selected articles.

The initial stage of topic identification was carried out by exploring various research databases and repositories to determine whether any previous reviews have addressed this specific topic. This initial search led to the formulation of the suggested research questions. The third stage included establishing a protocol for article selection. This stage is essential in outlining the search strategies and helps reduce error and bias. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) model was utilized in the selection of articles.

3.1. Literature search and selection procedure

PRISMA model serves as an evidence-based framework that is designed to assist academic writers in effectively reporting a wide range of systematic reviews and meta-analyses, typically utilized to assess the pros and cons of an intervention (Page et al., 2021). By employing PRISMA, the process of article selection for systematic reviews is rendered more transparent, consequently enhancing the overall quality of reporting within these reviews (PRISMA, 2025). Hence, the PRISMA Model was used in the current paper to screen the most relevant publications among the articles that were selected.

The primary search keywords utilized were “online language assessment”, “online language testing”, “attitude”, and “higher education”. Google Scholar, ERIC, and Scopus served as the primary databases employed for sourcing relevant papers. The initial search on Google Scholar included a time limit of 10 years and “online language assessment” as an exact search phrase, yielding 300 articles. The search on ERIC was restricted to peer-reviewed articles available in full text, resulting in 289 hits using the same keywords. The search on Scopus was performed using the specified keywords in addition to filters

including publication year (2015-2024), subject area (Arts and Humanities), and document type (Article), leading to 432 articles.

Fig. 1 represents the PRISMA flowchart in which the detailed process of article selection is illustrated. Before the screening process, 61 duplicate articles were identified and excluded. Then, 724 articles were excluded from the initial screening of the articles based on four criteria: (1) they did not focus on online language assessment or the English language, (2) they were not implemented in higher education, (3) were not published in peer reviewed journals, or (4) not written in English. In the second step of screening, articles were screened based on their titles and abstracts. Closer examination of abstracts and titles led to excluding 194 articles that were not relevant to the focus of the review or were not available in full text. Subsequently, 42 articles were eligible for full-text screening and 23 articles were excluded because (1) the article did not provide evidence of reliability or validity of methodology, (2) the article aim did not address the research questions of the review, (3) the article had missing information, or (4) the article was of poor quality or poor language. 19 articles are included in the review based on the selection made according to the four stages of the PRISMA model.

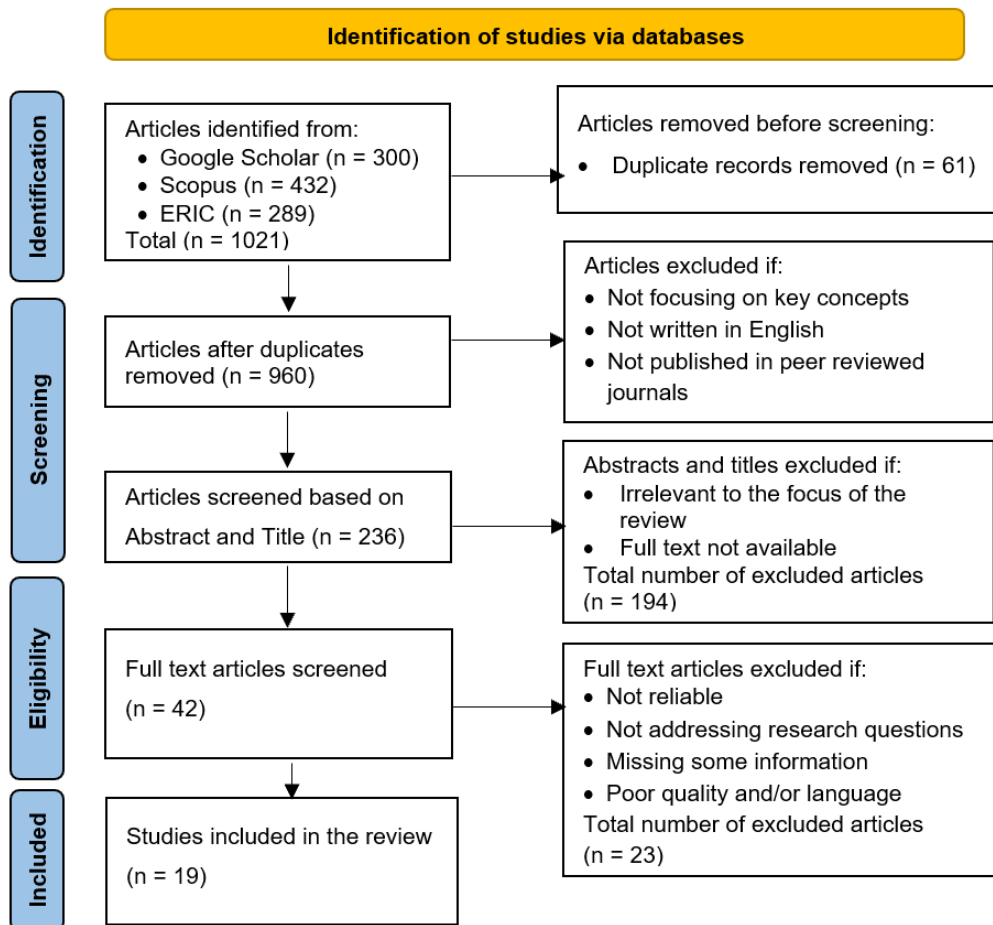


Fig. 1. PRISMA flowchart for literature search and selection

Table 2 demonstrates the articles that were found to address the research study questions. Typically, not all studies answer the three research questions; there are studies that answer only one question, while others answer two research questions. Only one study (Yang & Qian, 2017) was found to address the three questions of the current review.

Table 2

Overview of the included studies that address the research questions

No.	Studies	Question 1 (Effectiveness of online assessment)	Question 2 (Attitude of learners)	Question 3 (Challenges)
1.	Yang & Qian (2017)	√	√	√
2.	Elmahdi et al. (2018)		√	
3.	Mohamadi (2018)	√		
4.	Mir et al. (2019)		√	√
5.	Lu et al. (2020)	√	√	
6.	Zuhriyah & Pratolo (2020)		√	
7.	Alharbi & Meccawy (2020)		√	
8.	Sari et al. (2020)	√		√
9.	Yarahmadzahi & Goodarzi (2020)	√	√	
10.	Nurhadiani & Pratolo (2020)		√	
11.	Fitriyah & Jannah (2021)		√	√
12.	Ghanbari & Nowroozi (2021)			√
13.	Ghouali & Cecilia (2021)	√		
14.	Poonpon (2021)		√	√
15.	Alzubi et al. (2022)			√
16.	Mariadi et al. (2022)			√
17.	Zhang & Wu (2022)			√
18.	Ahmad et al. (2023)			√
19.	Koçer & Köksal (2024)			√

4. Findings and discussion

The current review study provides essential insights into the effectiveness of online language assessment, in addition to the attitudes of students towards it and the challenges it faces. This is implemented through reviewing several research studies that examine the relevant constructs as per the specified research questions. The present study introduces its findings by first providing relevant information about the included articles, such as methodology, participants, and key results, followed by a synthesis of the findings. The results of the reviewed studies are presented in relation to their alignment with the research questions.

Table 3 summarizes the methodologies and findings of the studies included in the review related to the study constructs. The research studies attempted to measure the effectiveness of the assessment tools by using pre-post-tests. By comparing online assessment to paper-based assessment, researchers sought to determine the benefits of online assessment. Several studies utilized research tools such as questionnaires, interviews, and observations to evaluate the attitude of students towards the use of different assessment platforms. Most of the studies used questionnaires, interviews, and observations to report on the challenges faced by learners and educators in relation to the use of online assessment.

Table 3
Summary of the included studies

No	Study	Data Collection	Instruments	Participants	Results
1.	Yang & Qian (2017)	MCQ test – questionnaire	• Reading comprehension tests • The Computer Dynamic Assessment (C-DA) software	Group 1 (52 students), Group 2 (46 students) freshmen majoring in industrial design at Lingnan Normal University in Mainland China	Effectiveness: The C-DA software outperformed traditional static assessments due to its strong diagnostic capabilities and deep insights. Attitude: C-DA was particularly helpful in: • building confidence for low achievers, • improve reading comprehension abilities. Challenges: They faced difficulties in: • understanding new vocabulary • locating relevant sentences • using inferencing abilities Attitude: The use of Plickers was found to: • support students' belief in the value of formative assessment and obtaining rapid feedback, • increase student engagement, • reduce instructional time, • ensure equitable participation chances, • foster an enjoyable and stimulating learning environment. Effectiveness: • The online collaborative writing evaluation intervention improved participants' writing skills, showing the highest level of significant writing enhancement. • The use of engaging technologies and approaches combined with sensible evaluation methods represents a potent method for enhancing learning effectiveness.
2.	Elmahdi et al. (2018)	A 17-item questionnaire with 3 open-ended questions	• Formative assessment • Plickers	166 students at the University of Bahrain	
3.	Mohamadi (2018)	Pre-/post-tests – Online portfolios	• Online summative, portfolio, formative and collaborative formative assessments • An electronic writing forum tool	130 Iranian (EFL) junior university students	
4.	Mir et al. (2019)	TAM survey on the M-assessment technology	• Formative assessment • Mobile SMS	101 BA & MA students	Attitude: • The M-assessment technology is practical, acceptable, fast, and cost-effective. • 87% of respondents were happy with the mobile assessment. Challenges: • Sometimes complex. • Has a high risk of plagiarism. • Requires training. Effectiveness: • There is a significant correlation between Peer Assessment (PA) and AES, supporting previous research. • Human raters and AES are consistent, with AES having higher validity.
5.	Lu et al. (2020)	A questionnaire	• Peer Assessment • AES system – online academic information service platform	46 EFL learners at a Chinese university	Attitude: Novice writers prefer AES over PA. Qualitative data analysis revealed new insights into PA in the EFL writing context.
6.	Zuhriyah & Pratolo (2020)	Semi-structured interviews	• Formative Assessment • Quizizz tool (Game-based)	6 students at a private university in Yogyakarta	Attitude: Quizizz was found to: • be an engaging tool, • help in boosting confidence,

7.	Alharbi & Meccawy (2020)	Pre-post surveys	• Formative Assessment • Socrative quizzes application	35 students at the Preparatory Year Program at a Saudi state university	• inspire more drive, • help students become better readers. The use of Quizizz in the classroom received favorable feedback from the students. Attitude: Initial survey results favoured paper-based tests over mobile tests, but following the experiment, the students' attitudes shifted in favour of mobile testing. Students prefer mobile-based tests for: • time saving • reduced anxiety • instant feedback • better translation • final score display • user-friendliness • and better focus due to one question per screen. Effectiveness: Shows benefits in terms of: • efficiency • conformity to students' characteristics • cost-effectiveness • detailed results • helpful features Challenges: • Limited features in scoring keys (e.g. short answer questions). • Inability to detect and/or control cheating. • Technical challenges include limited internet access or not having enough devices. Effectiveness: The treatment group that received continuous feedback from the application outperformed the control group. Attitude: The questionnaires revealed that they had positive opinions about employing mobile devices for testing.
8.	Sari et al. (2020)	Class activity log & observation checklist	• Formative Assessment • Google Forms	240 non-English major students at Sanata Dharma University in Indonesia	
9.	Yarahmadzhe i & Goodarzi (2020)	Pre/post-tests – treatment group that used the application and a control group that did not use the application – Attitude questionnaire	• Formative assessment • Socrative quizzes application	40 Pre-intermediate EFL learners	
10.	Nurhadianti & Pratolo (2020)	Semi-structured Interviews	• Formative Assessment • Kahoot! application	8 MA-level students at a private university in Yogyakarta	Attitude: • The application enhanced students' curiosities and elevated their motivation and learning objectives. • The application improved the quality of students' learning, with an emphasis on classroom environment, student engagement, and learning activity progress. • The application provides easy-to-use feedback to aid students in boosting their motivation to study. • The application may be used efficiently to gauge how well pupils understand the subjects they are being taught.
11.	Fitriyah & Jannah (2021)	A questionnaire & interviews	• Summative Assessment	30 Intensive Course Program students (questionnaire) & 3 EFL teachers, and 6 students (Interviews) at	Attitude: Students have a favourable attitude towards the delivery of online assessments, according to the questionnaire. There were four themes for positive effects as follows: • enhancing assessment flexibility, • improving evaluation versatility, • raising teachers' and students' awareness of online evaluations, • developing learner autonomy.

				the Islamic University in Kediri	Challenges: There are three themes for challenges, including: • difficulty of administering the evaluation, • decrease in teacher-student connection, • elevated test anxiety.
12.	Ghanbari & Nowroozi (2021)	In-depth semi-structured retrospective interviews & Teachers' narrative accounts	• Formative Assessment • A Learning Management System (LMS) & MOODLE	20 Iranian EFL teachers at Persian Gulf University	Challenges: Teachers experienced challenges in terms of: • technology (e.g., digital literacy, technology access, attitudinal factors) • academic dishonesty, • pedagogy (e.g., poor feedback mechanisms, absence of physical interaction), • affective dimension (e.g., stress management and social isolation), • administration (e.g., ambiguity of plans and policies). Effectiveness: • The treatment group outperformed the control group significantly. The e-assessment system in Moodle included educational, practical, and affective characteristics that contributed to students' improved results. • Moodle improved student performance by providing personalized feedback, remarks, and remedial solutions through tailored lessons and classroom explanations, and addressing individual paragraph correction and writing problems. Attitude: Test takers were satisfied with: • page layout, • page navigation or instructions, • forms for the writing part, • multiple choice item layout. Proctors were generally successful in performing their duties remotely, which goes in line with previous research. Challenges: Challenges identified were related to: • test takers' digital literacy, • experience in online testing, • students were not well prepared for the test, • weaker test takers may be tempted to cheat, • test anxiety and pressure.
13.	Ghouali & Cecilia (2021)	Pre-/post-tests using treatment & control groups	• Online Writing Formative Assessment • Moodle	42 first-year English Dept. EFL students at Tlemcen University (Algeria)	
14.	Poonpon (2021)	Questionnaire	• An online testing system • A high-stakes language proficiency test	218 Thai EFL university-level students	Challenges: Challenges identified were related to: • Online tests face the following challenges: • high possibility of cheating & impersonation, • lack of literacy in relation to technology, • lack of awareness of assistive assessment methods. Challenges: Teachers reported five main challenges with online assessments: • are difficult to construct, • require high technical skills, • take a long time to follow all the test procedures, • receive late submission, • allow for plagiarism.
15.	Alzubi et al. (2022)	A questionnaire & semi-structured interviews	Online language assessment	61 EFL teachers at Najran University, Saudi Arabia	
16.	Mariadi et al. (2022)	In-depth interviews & observations	Online language assessment	2 EFL Lecturers at a state university in Palembang	

17.	Zhang & Wu (2022)	Interviews and thematic analysis	Synchronous online learning and assessment	14 EFL university-level students in Southern China	Challenges: • Lack of a healthy learning/studying environment. • Developing learner autonomy. • Changing interaction patterns. • Adaptation to remote assessment.
18.	Ahmad et al (2023)	Structured written interview questions – follow-up oral interviews	Online language assessment tools	21 language instructors from 11 universities	Challenges: Interviews revealed ten challenges that form four categories of issues as follows: • ICT infrastructure, • lecturers' online assessment literacy, • language learners' commitment and integrity, • the need for specific online language assessment guidelines.
19.	Koçer & Köksal (2024)	Semi-structured interviews	Online Teaching and Assessment Tools	14 EFL University Lecturers	Challenges: The following challenges were reported: • Academic misconduct (plagiarism and cheating) • Difficulty in proctoring • Validity and reliability concerns

Fig. 2 shows the number of studies that addressed EFL online assessment effectiveness, attitude of learners, and the challenges that online assessment faces. The results show that more research can be carried out on EFL online assessment effectiveness to reach more conclusive results among different populations and using various online assessment tools.

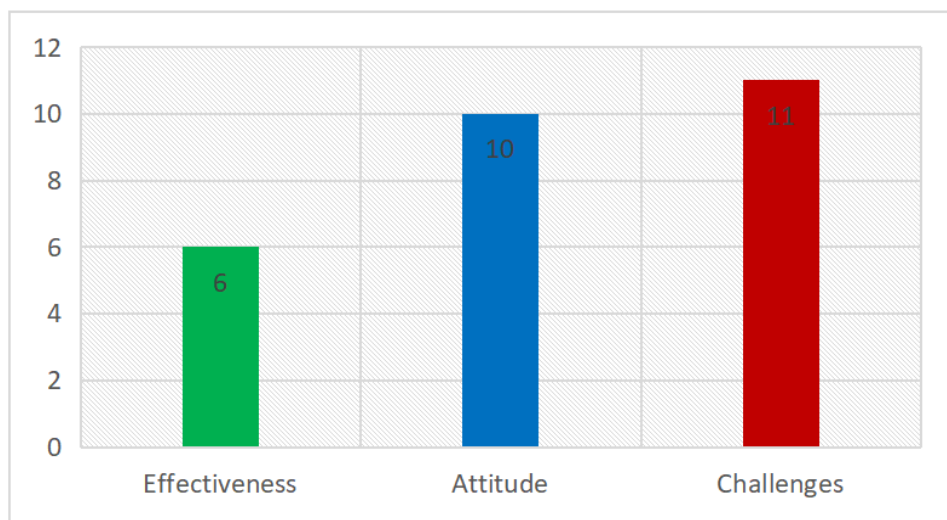


Fig. 2. Number of studies that addressed the effectiveness of online assessment, attitude of learners, and the challenges it faces

4.1. Effectiveness of online assessment in EFL education

The reviewed studies examined the effectiveness of using a range of educational platforms and applications and their ability to improve learners' abilities. All reviewed studies reported significant positive effects for the use of online assessment systems. Two studies

highlighted that the experimental groups that received online interventions from AES systems or collaborative writing evaluation interventions showed significant gains in writing performance due to the intensive feedback they received (Lu et al., 2020; Mohamadi, 2018). Lu et al. (2020) further reported that AES is more valid than PA in scoring essays and that novice teachers choose AES over PA.

Other studies that focused on the formative feedback that the learners received from Socrative Quizzes application, Google Forms quizzes, and Moodle showed significant performance improvement (Ghouali & Cecilia, 2021; Yang & Qian, 2017; Yarahmadzahi & Goodarzi, 2020). This significant improvement was reported to be the outcome of personalized feedback, detailed results and remarks, helpful features, and remedial solutions from formative feedback. Yang and Qian (2017) also found that the C-DA program outperformed traditional static assessments in assessing reading skills due to its stronger diagnostic capabilities and deeper insights.

The researchers of these studies concluded that online language assessment tools helped enhance learner engagement, promote equitable participation, and create a stimulating learning environment. These results suggest that EFL online assessment tools represent efficient tools in improving EFL learners' performance. However, more empirical research among different populations would yield more insights into the efficiency of other online tools and advanced AI-assisted online assessment systems.

4.2. Attitudes of learners towards EFL online assessment

The current review attempted to examine the attitudes of learners towards the use of online assessment. Among the reviewed studies, 10 reported on the attitudes and perspectives of learners towards the use of various online assessment tools such as C-DA, Plickers, M-assessment, Quizziz, Kahoot!, Socrative Quizzes application, and AES systems (Alharbi & Meccawy, 2020; Elmahdi et al., 2018; Fitriyah & Jannah, 2021; Lu et al., 2020; Mir et al., 2019; Nurhadiani & Pratolo, 2020; Poonpon, 2021; Yang & Qian, 2017; Yarahmadzahi & Goodarzi, 2020; Zuhriyah & Pratolo, 2020). All 10 studies reported positive feedback on the use of the various online tools in language assessment. This positive feedback in the studies can be discussed on three levels: affective level, pedagogical level, and technical level.

On the affective level, Yang & Qian (2017) highlighted that C-DA was particularly helpful in building confidence for low achievers, whereas the use of Plickers was found to increase student engagement and foster an enjoyable and stimulating learning environment (Elmahdi et al., 2018). Quizizz, on the other hand, was found to be an engaging tool, helps in boosting confidence and increases motivation. Learners also favoured mobile-based tests using the Socrative application for its reduced anxiety and had positive opinions about the application (Alharbi & Meccawy, 2020; Yarahmadzahi & Goodarzi, 2020). Kahoot! application was also reported to enhance students' curiosities and elevate their motivation (Nurhadiani & Pratolo, 2020). It also improved the quality of their learning, with an emphasis on classroom environment, student engagement, and learning activity progress. Learner positive comments on the feedback received via other online language assessment applications included raising online evaluation awareness.

On the pedagogical level, learners highlighted that C-DA was helpful in improving their reading comprehension abilities (Yang & Qian, 2017). Elmahdi et al. (2018) pointed out that the use of Plickers also had positive pedagogical gains, including reduced

instructional time and providing equitable participation chances. As for the Socrative Quizzes application, learners mentioned that it was helpful for time saving, instant feedback (Alharbi & Meccawy, 2020; Yarahmadzahi & Goodarzi, 2020). Kahoot! also improved the quality of learning as it can be efficiently used to gauge how well learners understand the content they are being taught and provide easy-to-use feedback that aids in boosting motivation (Nurhadianti & Pratolo, 2020). Learner further expressed their satisfaction with the use of other online assessment systems that enhanced assessment flexibility, improved evaluation versatility, and boosted learner autonomy (Alharbi & Meccawy, 2020; Fitriyah & Jannah, 2021).

On the technical level, users of the Socrative application praised its user-friendliness and reported better focus due to having one question per screen. The learners highlighted that Kahoot! was user-friendly and provided easy-to-use feedback (Nurhadianti & Pratolo, 2020). As for the Socrative Quizzes application, learners mentioned that it was helpful as it provided good translation and final score display options (Alharbi & Meccawy, 2020; Yarahmadzahi & Goodarzi, 2020). As for Google Forms, learners praised its detailed results and helpful features, while teachers mentioned that the application is cost-effective (Sari et al., 2020). Learners also expressed positive feedback on page and item layout and navigation options in another online assessment system (Poonpon, 2021).

These positive outcomes might inspire educators to consider online assessments as a viable alternative for evaluating language learners' performance formatively. Mir et al. (2019) supported this perspective, reporting high satisfaction rates among respondents with mobile assessment and M-assessment, emphasizing its efficiency in evaluating learner performance.

4.3. Challenges and limitations in EFL online assessment

Despite the positive attitude reported by students and the significant improvement in their performance, both teachers and learners expressed that they faced several challenges while using online language assessment tools in 11 studies (Ahmad et al., 2023; Alzubi et al., 2022; Fitriyah & Jannah, 2021; Ghanbari & Nowroozi, 2021; Koçer & Köksal, 2024; Mariadi et al., 2022; Mir et al., 2019; Poonpon, 2021; Sari et al., 2020; Yang & Qian, 2017; Zhang & Wu, 2022).

On the affective level, Fitriyah and Jannah (2021) highlighted that online assessment resulted in a decrease in teacher-student connection and elevated test anxiety. Ghanbari and Nowroozi (2021) also reported that teachers who used online assessment faced difficulties in stress management and suffered from social isolation, which comes in line with Fitriyah and Jannah's (2021) findings. Lack of a healthy learning/studying environment, anxiety due to changing interaction patterns, and difficulty adapting to remote assessment were also affective challenges reported by Zhang and Wu (2022).

On the pedagogical level, learners who used C-DA reported facing difficulty in understanding new vocabulary, locating relevant sentences, and using inferencing abilities while using the online system. Various studies reported a common challenge related to the high risk of plagiarism and ease of cheating and/or personification (Ahmad et al., 2023; Alzubi et al., 2022; Koçer & Köksal, 2024; Mariadi et al., 2022; Mir et al., 2019; Poonpon, 2021; Sari et al., 2020).

According to the reviewed studies, the technical level represents the biggest challenge in online language assessment. The users of M-assessment reported that the system is complex and requires training. Two other studies (Alzubi et al., 2022; Poonpon, 2021) highlighted that one major challenge is the lack of teachers' and students' digital awareness. This comes in line with the results reached by Yeşilyurt and Vezne (2023), who found that digital literacy and technological literacy significantly affected and explained the attitude of learners towards computer-assisted education. Another technical challenge was reported by Sari et al. (2020), who highlighted that Google Forms had poor feedback mechanisms. Other general technical challenges reported by teachers included limited internet access, limited features in scoring keys, not having enough devices, and difficulty with test construction.

Thus, online English language assessment has a few limitations and challenges that need to be addressed. Overcoming these challenges will boost the use of online language assessment tools over traditional paper-based assessments or even replace them, depending on the cultural, situational, financial, and political conditions. There is an increasing number of voices advocating for the enhancement of virtual education and the reduction of the digital divide, which is the gap between demographics and regions that have access to modern information and communication technology and those who have limited or restricted access (Saleh et al., 2023).

4.4. Recommendations and suggestions for future research

As reported in the current paper, online assessment in the EFL contexts has limitations and challenges that need to be addressed in order to boost the use of online language assessment tools. One major challenge reported in the current review was cheating and plagiarism, which are both forms of academic fraudulent behavior. Davidson and Coombe (2023) define other forms of fraudulent behavior in online assessment, such as impersonation (having someone take the exam in someone else's place), fabrication of data, and bribery or paid service. In their large-scale study (26,000 participants), Amorado and Vilchez (2023) found that 73% admitted to committing at least one form of fraudulent behavior in their online classes in various degrees and frequencies. These findings support the challenges reported in the current review.

With the increased demand for online assessment, a number of studies have been carried out to examine methods that can be used to minimize the cheating effect in online assessment (e.g., Amorado & Vilchez, 2023; Davidson & Coombe, 2023; Garg & Goel, 2023; Hadi, 2024; Lumpas et al., 2023; Mukherjee et al., 2023). Findings and recommendations of these studies suggest measures that can be taken to minimize probable fraudulent behavior. These measures can be classified into instructional measures, test design measures, and technical preparation measures.

Instructional measures that can be taken to minimize the cheating effect in online language assessment refer to the set of instructions that teachers/institutions give to their students before their exams in order to raise their awareness of the consequences and penalties of cheating. Various studies reported a significant effect for the use of an honor code reminder during an online test (Davidson & Coombe, 2023; Mukherjee et al., 2023) or signing an academic integrity contract with the learners (Hadi, 2024; Lumpas et al., 2023) in order to minimize the effect of cheating. Other instructional measures include teaching the students the various definitions of academic integrity and the different forms of academic fraudulent behavior, as well as the penalty for each.

Other research on academic integrity recommended measures that can be incorporated into the test design to minimize cheating. Lumpas et al (2023) highlighted that using questions that target higher-order thinking skills can significantly minimize the effect of cheating. Davidson and Coombe (2023) suggest other test design techniques, such as building an item bank for each course, which enables the teacher to assign random questions to students, or alternatively, using various versions of the exam. Other techniques include allowing only one question per page, setting a strict time limit for the exam, and adjusting the exam settings to auto-submit the exam after the time limit ends.

Recommendations on the technical arrangements before, during, and after online exams have been suggested by various researchers and test designers. For example, Davidson and Coombe (2023) recommend employing a process of multiple ID authentication before the exam in order to avoid impersonation. They also recommend the use of a secure browser, which is also recommended by Garg and Goel (2023), who reported that the use of tab-switch functionality – which is the ability of a browser to detect and control the number of opened tabs – helped in minimizing the cheating effect significantly. Although live proctoring was recommended by Lumpas et al (2023) to mitigate against cheating in online exams, Davidson and Coombe claimed that it is unsuccessful with large numbers and recommended the use of online proctoring companies such as Examity, Honorlock, ProctorCam, Respondus, etc. Davidson and Coombe added that the best proctoring technique is advanced automated proctoring in which an AI system records the test takers in real time and detects unusual actions they make. The AI system sends an instant warning to the teacher to report on possible cheating incidents that can then be closely examined by the human proctor.

The use of advanced AI systems in online assessment and proctoring is still a recent field of study. It is recommended that more research on the use of AI proctoring systems should be carried out in order to examine their effectiveness and how students perceive their use. More research is also recommended on the anxiety associated with online language learning in higher education.

Author Statement

The authors declare that there is no conflict of interest.

ORCID

Amr Selim Wannas  <https://orcid.org/0000-0001-8990-4943>

Marwa M. AbdelMohsen  <https://orcid.org/0009-0004-6902-9706>

References

- Acar, A., & Kayaoğlu, M. (2020). Moodle as a potential tool for language education under the shadow of COVID-19. *Eurasian Journal of Educational Research*, 90, 67–82. Retrieved from <https://ejer.com.tr/moodle-as-a-potential-tool-for-language-education-under-the-shadow-of-covid-19/>
- Ahmad, N., Ab Rahim, I. S., & Ihsan, P. I. (2023, May). Overview of challenges in conducting English for specific academic purposes (ESAP) course assessments via

- online and distance learning mode (ODL) for Engineering students amidst Covid-19 pandemic. In *Proceedings of the 1st UM Surabaya Multidisciplinary International Conference 2021 (MIcon 2021)* (pp. 369–381). Atlantis Press. https://doi.org/10.2991/978-2-38476-022-0_40
- Alderson, J. C., Brunfaut, T., & Harding, L. (2017). Bridging assessment and learning: A view from second and foreign language assessment. *Assessment in Education: Principles, Policy & Practice*, 24(3), 379–387. <https://doi.org/10.1080/0969594X.2017.1331201>
- Alharbi, A. S., & Meccawy, Z. (2020). Introducing Socrative as a tool for formative assessment in Saudi EFL classrooms. *Arab World English Journal*, 11(3), 372–284. <https://doi.org/10.2139/ssrn.3705759>
- Alzubi, A. A. F., Al-Mwzaiji, K. N. A., & Nazim, M. (2022). Online and offline assessment methods in higher education: A revisitation of EFL teachers' perceptions and practices. *Journal of Language Teaching and Research*, 13(6), 1147–1155. <https://doi.org/10.17507/jltr.1306.02>
- Amorado, R. V., & Vilchez, R. N. (2023). Remote integrity and accountability: Baseline profile of academic fraud and cheating in online classes. *Migration Letters*, 20(S11), 433–477. Retrieved from <https://migrationletters.com/index.php/ml/article/view/5662/3873>
- Bai, J. Y. H., Zawacki-Richter, O., Bozkurt, A., Lee, K., Fanguy, M., Cefa Sari, B., & Marín, V. (2022, September). Automated essay scoring (AES) systems: Opportunities and challenges for open and distance education. In *Proceedings of the Tenth Pan-Commonwealth Forum on Open Learning (PCF10)*. Commonwealth of Learning. <https://doi.org/10.56059/pcf10.8339>
- Chun, D., Kern, R., & Smith, B. (2016). Technology in language use, language teaching, and language learning. *The Modern Language Journal*, 100(S1), 64–80. <https://doi.org/10.1111/modl.12302>
- Crisp, G. (2011). *Teacher's handbook on e-assessment*. Australian Learning and Teaching Council. Retrieved from http://transformingassessment.com/sites/default/files/files/Handbook_for_teachers.pdf
- Davidson, P., & Coombe, C. (2023). Maintaining the integrity of assessment by mitigating against students' cheating. In I. Hande et al. (Eds.), *Re-Vision and Re-Form in English for Academic Purposes Contexts after the Pandemic* (pp. 66–74). Cambridge Scholars Publishing. Retrieved from <https://zuscholars.zu.ac.ac/works/6401/>
- Ebadi, S., & Saeedian, A. (2016). Planning future instructional programs through computerized L2 dynamic assessment. *Teaching English with Technology*, 16(4), 12–32. Retrieved from <https://tewtjournal.org/download/3-planning-future-instructional-programs-through-computerized-l2-dynamic-assessment-by-saman-ebadi-and-abdulbaset-saeedian/>
- Elmahdi, I., Al-Hattami, A., & Fawzi, H. (2018). Using technology for formative assessment to improve students' learning. *TOJET: Turkish Online Journal of Educational Technology*, 17(2), 182–188. Retrieved from <https://eric.ed.gov/?id=EJ1176157>
- 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>
- Garg, M., & Goel, A. (2023, January). Detection of internet cheating in online assessments using cluster analysis. In *Proceedings of the Seventh International Conference on Data Management, Analytics & Innovation (ICDMAI 2023)* (pp. 77–90). Springer.

- https://doi.org/10.1007/978-981-99-1414-2_7
- Ghouali, K., & Cecilia, R. R. (2021). Towards a Moodle-based assessment of Algerian EFL students' writing performance. *Porta Linguarum*, 2021(36), 231–248. Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=8020384>
- Ghanbari, N., & Nowroozi, S. (2021). The practice of online assessment in an EFL context amidst COVID-19 pandemic: Views from teachers. *Language Testing in Asia*, 11: 27. <https://doi.org/10.1186/s40468-021-00143-4>
- Granić, A. (2023). Educational technology adoption: A systematic review. *Education and Information Technologies*, 27(7), 9725–9744. <https://doi.org/10.1007/s10639-022-10951-7>
- Hadi, K. (2024). Some teaching tips for overcoming challenges and conducting an effective online assessment in Algerian universities. *Journal of Languages and Translation*, 4(1), 66–79. Retrieved from <https://journals.univ-chlef.dz/index.php/jlt/article/view/585>
- Hadijah, H., Pratolo, B. W., & Rondiyah, R. (2020). Interactive game “Kahoot!” as the media of students' vocabulary assessment. *Journal on English as a Foreign Language*, 10(1), 84–102. <https://doi.org/10.23971/jefl.v10i1.1670>
- Hussein, M. A., Hassan, H., & Nassef, M. (2019). Automated language essay scoring systems: A literature review. *PeerJ Computer Science*, 5: e208. <https://doi.org/10.7717/peerj-cs.208>
- Jurāne-Brēmāne, A. (2023). Digital assessment in technology-enriched education: Thematic review. *Education Sciences*, 13(5): 522. <https://doi.org/10.3390/educsci13050522>
- Kamrood, A. M., Davoudi, M., Ghaniabadi, S., & Amirian, S. M. R. (2019). Diagnosing L2 learners' development through online computerized dynamic assessment. *Computer Assisted Language Learning*, 34(7), 868–897. <https://doi.org/10.1080/09588221.2019.1645181>
- Koçer, P., & Kökeh, D. (2024). An investigation into the online language teaching and assessment practices during COVID-19. *International Journal of Educational Spectrum*, 6(1), 1–17. <https://doi.org/10.47806/ijesacademic.1341237>
- Kolitsnyk, M., Kornitska, Y., Ogurtsova, O., & Sokyrskya, O. (2022). Socrative as a formative assessment tool in English for specific purposes course. *Arab World English Journal*, 2nd Special Issue, 377–392. <https://doi.org/10.24093/awej/covid2.25>
- Lu, M., Deng, Q., & Yang, M. (2020, February). EFL writing assessment: Peer assessment vs. automated essay scoring. In *Proceedings of the 4th International Symposium on Emerging Technologies for Education (SETE 2019)* (pp. 21–29). Springer. https://doi.org/10.1007/978-3-030-38778-5_3
- Lumpas, C. M., Wagia, J., Balayanan, J., & Acelar, S. J. (2023). Discovering teachers' strategies to mitigate online cheating: An exploratory design. *Southeast Asian Journal of Multidisciplinary Studies*, 3(1).
- Maret'atush S, K., Nur Afifah, W. F., Indriani, N., Ini'matul R, A., & Oktovianti D, N. F. (2021, June). The effectiveness of google form as a medium for assessing students' english learning outcome in high school level. In *Proceeding of 1st International Conference on Enhancing English Language Teaching through the Use of Technology in the Industry 4.0 Era (CELTi)* (pp. 112–124). Retrieved from <https://proceedings.uinsaizu.ac.id/index.php/celti/article/view/399/647>
- Mariadi, M., Mirizon, S., & Sitinjak, M. D. (2022). The challenges of online assessment amidst covid-19 pandemic disruption: The voice of EFL lecturers. *Englisia: Journal of Language, Education, and Humanities*, 10(1), 107–126. <https://doi.org/10.22373/ej.v10i1.13704>
- Mir, K., Iqbal, M. Z., & Shams, J. A. (2019). An investigation of AIOU students'

- satisfaction about formative M-Assessment using SMS technology. *Pakistan Journal of Distance and Online Learning*, 5(2), 157–174. Retrieved from <https://eric.ed.gov/?id=EJ1266716>
- Mohamadi, Z. (2018). Comparative effect of online summative and formative assessment on EFL student writing ability. *Studies in Educational Evaluation*, 59, 29–40. <https://doi.org/10.1016/j.stueduc.2018.02.003>
- Mukherjee, S., Rohles, B., Distler, V., Lenzini, G., & Koenig, V. (2023). The effects of privacy-non-invasive interventions on cheating prevention and user experience in unproctored online assessments: An empirical study. *Computers & Education*, 207: 104925. <https://doi.org/10.1016/j.compedu.2023.104925>
- Nurhadianti, B. W. P. (2020). Students' perception toward the application of Kahoot! as an assessment tool in EFL class. *Universal Journal of Educational Research*, 8(5), 2150–2155. <https://doi.org/10.13189/ujer.2020.080554>
- Örnfelt, J. (2023). *Users' attitude towards ChatGPT: A sentiment analysis on Twitter & Reddit*. Master dissertation, Uppsala University, Sweden. Retrieved from <https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-505786>
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *British Medical Journal*, 372: 160. <https://doi.org/10.1136/bmj.n160>
- Poonpon, K. (2021). Test takers' perceptions of design and implementation of an online language testing system at a Thai university during the COVID-19 pandemic. *PASAA*, 62(1), 1–28. <https://doi.org/10.58837/CHULA.PASAA.62.1.1>
- PRISMA. (2025). *PRISMA flow diagram*. Retrieved from <https://www.prisma-statement.org/prisma-2020-flow-diagram>
- Purpura, J. E. (2016). Second and foreign language assessment. *The Modern Language Journal*, 100(S1), 190–208. <https://doi.org/10.1111/modl.12308>
- Radhaswati, I. D. D. A., & Santosa, M. H. (2022). Teachers' perceptions: The use of google forms as a media to assess primary school students. *EDUTEC: Journal of Education and Technology*, 5(4), 910–924.
- Saleh, M. A., Khan, M. I., Banerjee, S., & Safi, F. (2023). A tale of online learning during COVID-19: A reflection from the south Asian association for regional cooperation (SAARC) countries. *Heliyon*, 9(6): e16347. <https://doi.org/10.1016/j.heliyon.2023.e16347>
- Sari, A. B. P., Iswahyuni, D., Rejeki, S., & Sutanto, S. (2020). Google forms as an EFL assessment tool: Positive features and limitations. *Premise: Journal of English Education and Applied Linguistics*, 9(2), 231–250. <http://dx.doi.org/10.24127/pj.v9i2.3037>
- Schweighofer, J., Taraghi, B., & Ebner, M. (2019). Development of a quiz-implementation of a (self-) assessment tool and its integration in Moodle. *International Journal of Emerging Technologies in Learning*, 14(23), 141–151. <https://doi.org/10.3991/ijet.v14i23.11484>
- Shadiev, R., & Yang, M. (2020). Review of studies on technology-enhanced language learning and teaching. *Sustainability*, 12(2): 524. <https://doi.org/10.3390/su12020524>
- Sievertsen, H. H., & Burgess, S. (2020). *Schools, skills, and learning: The impact of COVID-19 on education*. Retrieved from <https://cepr.org/voxeu/columns/schools-skills-and-learning-impact-covid-19-education>
- Titova, S. (2019). Mobile testing system for developing language skills. In M. Khosrow-Pour (Ed.), *Encyclopedia of Information Science and Technology* (pp. 5116–5126). IGI Global. <https://doi.org/10.4018/978-1-5225-2255-3.ch444>

- Wang, A. I., & Tahir, R. (2020). The effect of using Kahoot! for learning – A literature review. *Computers & Education*, 149: 103818. <https://doi.org/10.1016/j.compedu.2020.103818>
- Wong, W. S., & Bong, C. H. (2019). A study for the development of automated essay scoring (AES) in Malaysian English test environment. *International Journal of Innovative Computing*, 9(1), 69–78. <https://doi.org/10.11113/IJIC.V9N1.220>
- Yang, Y., & Qian, D. D. (2017). Assessing English reading comprehension by Chinese EFL learners in computerized dynamic assessment. *Language Testing in Asia*, 7(1): 11. <https://doi.org/10.1186/s40468-017-0042-3>
- Yarahmadzahi, N., & Goodarzi, M. (2020). Investigating the role of formative mobile based assessment in vocabulary learning of pre-intermediate EFL learners in comparison with paper-based assessment. *Turkish Online Journal of Distance Education*, 21(1), 181–196. <https://doi.org/10.17718/tojde.690390>
- Yeşilyurt, E., & Vezne, R. (2023). Digital literacy, technological literacy, and internet literacy as predictors of attitude toward applying computer-supported education. *Education and Information Technologies*, 28(8), 9885–9911. <https://doi.org/10.1007/s10639-022-11311-1>
- Zhang, K., & Wu, H. (2022). Synchronous online learning during COVID-19: Chinese university EFL students' perspectives. *Sage Open*, 2022(4), 1–10. <https://doi.org/10.1177/21582440221094821>
- Zuhriyah, S., & Pratolo, B. W. (2020). Exploring students' views on the use of quizizz as an assessment tool in English as a foreign language (EFL) class. *Universal Journal of Educational Research*, 8(11), 5312–5317. <https://doi.org/10.13189/ujer.2020.081132>