

Research Article

Evaluating Assessment Strategies for EFL Teacher Pre-, During-, and Post-COVID-19: Teachers' Reflections on Challenges and Opportunities

Hassan Alizadehmahmoudalilo

Assistant Professor, Department of English Language, Ta.C., Islamic Azad University, Tabriz, Iran,
hassan.alizadehmahmoudalilo@iau.ac.ir

ARTICLE INFO

Submission History

Received: 2025-12-02
Accepted: 2025-12-30

Keywords

Assessment Literacy
COVID-19
EFL Teachers
Formative Assessment
Summative Assessment

ABSTRACT

The COVID-19 pandemic has transformed education globally in an unprecedented way; therefore, it has become critical to reexamine how we assess students, particularly when it comes to the teaching of languages. The current study examines the pre-COVID-19 situation, during COVID-19, and post-COVID-19, and the change in assessment by the Iranian EFL teachers. The current state of research is minimal as to the assessment practice in postsecondary education; therefore, it resorts to a qualitative, phenomenological design in order to study what has been practiced. The 25 experienced teachers provided data using semi-structured interviews and open-ended questionnaires. Thematic analysis demonstrated that the pre-pandemic evaluation was mostly summative-based. The postgraduate courses, on the contrary, permitted a more formative and research-oriented approach. Initially, the pandemic forced the teachers to revise the traditional use of tests, but with time, they began relying on blended models to address issues such as dishonesty, lack of technology, and students who were not interested in school. Following the pandemic, blended and formative techniques were still applied by teachers, yet evaluating how to maintain innovation and equity in assessment was of their concern as well. Respondents reported that they are better equipped with information on the use of technology and developing new methods of teaching it; however, they experienced trouble with the validity of assessments and the condition of the infrastructure. As indicated by the findings, we require the concept of language assessment literacy, equitable online infrastructure, and continuous professional growth to foster assessing the subjects through a learner-centered approach.



Introduction

The COVID-19 pandemic significantly affected the educational settings all over the world, resulting in an almost overnight transition to online teaching. This shift in Iran had a great impact on teaching and assessment in the EFL programs. Assessment is a crucial EFL aspect because composing, speaking, and reading critically demand evaluation approaches that are specific and reliable (Pokhrel & Chhetri, 2021). Although a lot of research focused on online learning during the pandemic, there is less evidence on assessment adjustments concerning EFL programs in Iran.

Assessment is a major element of proper teaching and learning through providing important insights into student attitudes as well as behavior, informing pedagogical choices, and helping in promoting academic development (Black & Wiliam, 2009). Teaching English as a Foreign Language (TEFL) is an area of teaching in which such advanced skills as academic writing, critical reading, and oral communication have to be tested, and the online assessment system presented serious challenges. The sudden transformation forced the educators to re-evaluate and, in most cases, restructure their assessment plans for the digital settings.

Recently, numerous studies have been conducted to illuminate this change. As an example, Olasina (2023) examined the implementation of new digital assessment applications in the course and after the pandemic, which became increasingly relevant in education after COVID-19. These tools provided flexibility and expediency, but triggered problems associated with equity, validity, and preparedness of the digital environment. The difficulties regarding digital literacy levels, technology access, and support

by the institution, especially to the ESL lecturers who change to the new technologies, were raised by Emelogu et al. as well (2022). Such results applaud the Iranian higher education environment, where digital infrastructure and pedagogical education are unevenly spread out.

As mentioned earlier, COVID-19 has affected education in general in broad terms; however, few people have examined its scope in the field of language assessment, specifically, on EFL teachers. Although numerous researchers investigated the motivation and online learning attitudes of learners (Bozkurt & Sharma, 2020), not much attention was given to the assessment strategy that was developed throughout the course and could be perceived by authorities. Another example, Hunutlu et al. (2023) conducted a study on language assessment literacy (LAL) of EFL instructors in Turkey and revealed that, along with the enhancement of digital knowledge, it was challenging to pedagogically align the assessment objectives. In their research, they emphasize the necessity to improve the knowledge regarding which tools were employed and how efficient they were in specific assessment conditions.

The abovementioned complexities can be observed in the context of Iran as well. Momeni (2022) studied the opinions of Iranian EFL teachers towards online assessment during the pandemic, and the research indicated various attitudes: some liked the level of flexibility offered by remote tools, but most teachers encountered the problem of fairness, engagement, and authenticity. Referring to the opportunities of online writing assessment during COVID-19, Al-Bargi (2022), in a similar context to Iran, found a similar conflict between the usefulness

of online resources and the fear of plagiarism and low-level student engagement.

Besides these logistics and technological issues, the pandemic has also compelled a more critical approach to how and why assessments are designed by the educators. Tian et al. (2021) envisioned COVID-19 as a critical incident that exposed the flaws of the current state of language assessment strategy and demanded a revolutionary reconsideration of AL. The article by Ashrafian and Alimohammadi (2024) also focused on the targeted professional development regarding digital assessment practice among EFL instructors in post-pandemic Iran and posed a suggestion that educators identified the changes in their roles, yet they do not find adequate support within their institutions.

Along with that, the dual nature of digital tools is called upon in the literature as well, granting flexibility, and on the other hand, posing the question of validity, digital equity, and instructor readiness (e.g., Emelogu et al., 2022; Olasina, 2023). LAL has become of major concern by the instructors, and the balance of such evaluation and the degree to which summative assessment (SA) can be more informative than formative assessment (FA) is now less rigid in the virtual world (Nicol, 2007). However, the majority of the teachers were capable of adapting to creative solutions, but there is little research that can evaluate these adaptations in the TEFL courses. It is within this background that the current research attempts to address a significant gap in the body of literature by conducting an exploration of the experiences of Iranian EFL teachers, in measuring the TEFL learners regarding the impact and angle of pre- and post- as well as during the COVID-19 pandemic.

Literature Review

This research discusses the way the evaluation methods of EFL teachers have changed, as they have depended on a conventional and summative approach to digital and blended assessment, particularly because of the COVID-19 outbreak. In this part, world literature and concepts in Iran are referred to in examining the problem of SA vs. FA and theories of approach (e.g., AL, learning-oriented assessment, constructivist approaches) and the pedagogical role of assessment in the language advanced learning process. Consider then, the pre-pandemic practices, the COVID-19 disruptions, online/blended formative learning and teacher AL and perception, technological obstacles, post-pandemic trends, and research gaps.

1. Conceptual Foundations

SA in TEFL can be understood as an end-of-course assessment with marks or grades, and FA implies feedback throughout the course to aid the process of learning. Students, as an example, used to be given limited feedback in the form of a final score, which is provided by SA and is used to assess learning outcomes (Glazer, 2014). Conversely, FA enables feedback that enhances learning and entails the use of assessments as frequent, interactive to identify the needs of students and make necessary adjustments in teaching (Ismail et al., 2020). FA becomes a part of the learning experience (quizzes, drafts, peer review, etc.), but in SA, it is usually summed up by the high-stakes exams or final projects. A combination of summative and formative approaches is proposed by some researchers, e.g., assessment of examinations followed by the provision of feedback (Alahmadi et al., 2019; Remmi & Hashim, 2021).

AL can be described as the knowledge and expertise of teachers in the creation, administration of assessments, and assessment interpretation (Weng & Shen, 2022). It is the key to successful instruction (Popham, 2009), yet it needs to be developed more often than not (Berry et al., 2017). The competence of language teachers in designing valid tests and feedback loops is called LAL; the absence of LAL may contribute to the misdesign of the assessment and flawed educational choices (Weng & Shen, 2022). Learning-Oriented Assessment (LOA) is a new method that integrates assessment into learning: it lays an emphasis on learning and is aimed at involving learners into the assessment and feedback practices (Al Kamli & Almalki, 2024). LOA does not just measure the standard assessment but looks at it as a process combined with teaching (Davison, 2019). As an example, Al Kamli and Almalki (2024) mention that in LOA, assessment is intended to facilitate and enhance learning; the assessment is carried out using formative quizzes, self- and peer-assessment, as well as reflective activities. Assessment is also affected by constructivist theories (Piaget, Vygotsky): these theories focus on learners to construct knowledge and argue that assessment needs to focus on active, social and contextualised tasks (e.g. tasks or projects with feedback) as opposed to rote testing (Vogt & Tsagari, 2024). In constructivism, formative assessment is dialogic since it shifts the notion of learning as co-constructed (Liu et al., 2021).

Assessment has its special functions at the postgraduate level. It promotes higher-level academic writing and research abilities, as well as professional competencies. An example is formative assessment (in the form of feedback on drafts of theses) of helping master students to write in a scholarly manner. Students (as

Esfandiari et al., 2022, also revealed in the case of Iranian EFL postgraduates) want to be heard, engaging their needs (e.g., writing support), and having consistent evaluation standards are crucial, as students find their requirements overlooked and face the lack of effective assessment based on stable criteria. AL between TEFL teachers is therefore essential. They should be capable of creating valid tasks (research proposals, literature review, etc.) and feedback that would develop research literacy and autonomy among the teachers (Sohrabi et al., 2022). In assessments within graduate TEFL education, the development of the profession: self-evaluation of professional preparedness (i.e., by undertaking student assessment (e.g., in microteaching, portfolio compilation)) helps to promote reflection and get graduate students ready to teach. To conclude, postgraduate language program assessment must be a process of certification and at the same time an opportunity to actively promote advanced learning, writing proficiency, and scholarly skills (Weigle, 2007).

2. Pre-COVID Assessment Practices in TEFL

Until recently, the programs of TEFL obtained in postgraduate classes were mostly based on presenting conventional assessment forms. They usually focused on summative or exam-type assessments. In most Asian and even some Middle-Eastern EFL settings, language courses (at least those aiming at undergraduate degrees) focused on grammar/vocabulary tests, regular tests on writing, and final projects as ways to conclude the course (Vellanki et al., 2023). To give an example, in their review, Vellanki et al. (2023) observe that the usage of in-class quizzes, writing activities, and portfolios is usually coupled with periodical, written examinations

or projects in educational institutions of higher learning. In those contexts, some formative activities did take place (e.g., oral presentations, peer review), but the benchmark that mattered most was the final exam or the term paper. This kind of summative supremacy was indicative of educational culture; a recent study on the EFL language class context gave evidence that standards on exam scores were enormously significant; it put instructors under societal responsibilities to avert cheating under these high-stakes tests (Vellanki et al., 2023).

Studies that refer to exam-based assessment in Iran and similar EFL frameworks (Turkey, East Asia, Middle East) tend to characterize it. Traditional English programs have employed paper and pencil, teacher-made exams (listening, speaking, and writing skills) that were conducted in controlled testing centers. According to Crusan et al. (2016), professionals in the curriculum field rarely review teacher assessment practice in the field, which results in the continuation of a test-oriented culture. In the case of postgraduate, though, assignments and presentations have received some attention. According to Sohrabi et al. (2022), writing students at Iranian MS took part in inappropriately introduced and inconsistently evaluated courses, which means that the level of formative process involvement should be increased. In much the same vein, in South Korea, Lee (2017) observed that, thus far, grad school English programs tend to have an evaluation of research projects and capstone work, but standardized tests (such as English proficiency) are still used. However, a consistent pre-COVID criticized issue is that even higher-level testing is based too frequently on measured results and not learning procedures (Momeni, 2022).

In TEFL, pre-pandemic assessment has been criticized for serving as test-driven learning and providing little feedback. The value of traditional summative tests is limited to formative assessment, and educators occasionally do not use exams as a resource (Glazer, 2014). Remmi and Hashim (2021) have noted that most educators regard the FA as a simple metric tool and fail to realize its instructional value. In addition, graduate TESOL programs do not usually offer courses in language assessment, and so, new educators join the margins of the field without adequate LAL (Weigle, 2007). Such variance can be observed in Iran: as Ashrafian and Alimohammadi (2024) state, the number of EFL lecturers with little to no training in innovative assessment practices is large in Iran (assessment knowledge was identified as the highest professional development priority among teachers). Therefore, the pre-COVID TEFL evaluation was inclined towards the summative focus, and the postgraduate learners were often forced to receive limited formative feedback (e.g., final thesis-oriented feedback only), which could be improved by involving assessment to support the development of higher-level skills.

3. Impact of COVID-19 on Assessment

The COVID-19 pandemic was a global intrusion in education systems that has never been seen before (Ali et al., 2020). It compelled an extreme shift from face-to-face testing and evaluation to distant assessment. At the international level, schools and universities rushed to meet the lockdown requirements: numerous tests were canceled, pushed to a dead date, or turned into remote forms (Vogt & Tzagari, 2024). In particular, there was the postponement or switching of high-stakes exams to former evaluations in some

universities (Maaoui et al., 2023). Language tests in all of Canada were canceled. In other states, they moved to online opportunities; Ngololo (2021) cites universities that launched remote proctored exams. Still, they were greeted by backlash in parts: in Kenya, South Africa, and Zimbabwe, protests against compulsory online testing started, with most students afraid of making disparities greater (Kigotho, 2020).

To proceed with assessment, EFL educators resorted to all kinds of digital tools. Oral exams and speaking activities took place in video channels such as Zoom, Google Meet, or Microsoft Teams. Quizzes, assignments, and moderation of discussion topics in Learning Management Systems (LMSs) (e.g. Moodle, Google Classroom, Blackboard) were utilized. Social media and web forms were also utilized: an Iranian study mentioned that Google Forms and Classroom have received widespread use in quizzes and submissions (Momeni, 2022). Online remote proctoring suites were used by test providers for proctored foreign language exams (Purpura et al., 2021). In the meantime, asynchronous tools (e-portfolios, recorded presentations, project-based tasks) grew instead of the conventional in-person tests (Vogt & Tsagari, 2024). In particular, the Turkish teachers embraced the take-home test, open-ended assignments, and project work as the best tests to be delivered remotely (Şenel & Şenel, 2021). Therefore, the pandemic hastened the process of establishing various ways of e-assessment in TEFL.

There were some benefits that were brought by digital assessment. Flexibility was enhanced, and students could do their tests at home or access feedback online. Such tests as synchronous online speaking tests allowed discussions at the level of one or more people

without any traveling. Video projects, discussion forums (i.e., asynchronous methods) involved the learners creatively. In the study of Sutadji et al. (2021), a combination of written case studies, online debates, and peer review produced a more natural form of assessment that best corresponded to the usage of the language. As noted by Vogt and Tsagari (2024), one of the outcomes of emergency remote assessment was the fast development of innovative types of assessments suggested by test developers who created online assessment opportunities through online test development while maintaining security (e.g., proctoring, test adaptability). More broadly, the crisis led to the initial proliferation in numerous universities of tools (e.g., LMS, videoconference), which previously existed merely as supplements. These changes introduced an opportunity for being creative and sustainable: the teachers organized learning communities to discuss formative and multimodal tasks, which made students more engaged in their activities (Chung & Choi, 2021). The usage of new interactive forms (forums, projects) was usually appreciated by students in surveys.

Nevertheless, remote assessment is accompanied by a variety of challenges. Ethical integrity in education has gained top priority in the world. When exams were done without supervision, the chances of cheating also increased; the teachers decided that the students were employed during the examinations using search engines or WhatsApp groups (Momeni, 2022). Online proctoring promised security and brought out concerns of privacy and fraud. It was a big challenge in terms of digital access and inequality. In developing countries, fair testing was prevented by unreliable internet and a

shortage of devices. According to one of the Iranian research studies (Momeni, 2022), poor internet was mentioned by a quarter of teachers as one of the main difficulties, and inadequate technology infrastructure was cited by 18% of all teachers. The interviews showed stress as students left Zoom exams to cheat on the exams of others. The digital divide was clear across the globe: in Africa, demonstrations were voiced against the fact that online assessment widened the gap between urban and rural communities. Teacher readiness was another issue. Most of the instructors were uneducated or weak on online evaluation. In a survey conducted by Vellanki et al. (2023), the teachers were forced to go by trial and error, where the first approach they took was one they knew, most commonly an open-book exam. Problem with technology (LMI problems, did not know LMS interfaces) frequently occurred. According to the study performed by Momeni (2022), the lack of technological know-how is one of the major barriers perceived by teachers. Although the platforms were present, the developed ones were considered as supplements by many, and people were not ready to take a fully online course (Ghanbari & Nowroozi, 2021). Therefore, the problem of the lack of teacher digital literacy retarded implementation.

Overall, the crisis of COVID-19 transformed conventional TEFL assessment radically: institutions and teachers quickly used online coursework (LMS, video conferencing, automated quizzes) to maintain traditional evaluation and found empowering it (flexibility, real-life tasks) and acute disempowering (cheating, connectivity problems, digital literacy drawbacks).

4. Shift Toward Blended and Formative Approaches

Innovations that were forced due to the pandemic promoted an increasing tendency toward blended instruction and formative-based assessment in TEFL. Due to the transition back to school (or to hybrids), numerous schools have not lost online aspects and formative practice found during COVID. According to researchers, FA techniques are utilized more and more by teachers in their digital classrooms. In another example, Chung and Choi (2021) identified that new methods of assessments used by instructors had become formative and process-based when it came to online teaching. As an illustration, teachers did not use timed tests, but provided students (and peers) with iterative tasks (peer reviews, drafts, portfolios), which were to be continually evaluated. Combining several assessment methods, such as case-based writing tasks, student-led online discussions with peer feedback, and active learning tracking, makes evaluation more authentic, as reported by Sutadji et al. (2021). The community of professional development was created in which teachers were empowered to formulate collaborative projects and multimodal products instead of multiple-choice assessments (Isbell et al., 2023).

A combination of project-based learning and peer feedback have been mentioned to be effective in postgraduate EFL. Project based learning can involve the learners in a purposeful work in the real world (e.g. research project, presentation where the use of language has a purpose (Al Kamli & Almalki, 2024). Migration to FA and blended models have paid off, particularly on the areas of engagement and learner confidence. The students are less nervous during their performance, owing to the minimization of the

use of final examinations; they also get time to do more to correct their grades before they are given out by getting little reports. Importantly, FA enhances critical thinking; educators may present more complex, open-ended work in which they might ask students to analyze research articles, create lesson plans, and revise them several times (Ismail et al., 2022). Şenel and Şenel (2021) also identified an increase in ownership and satisfaction in the use of at-home open-ended projects in Turkey during COVID because students were able to make sense of language. Overall, the combination of online tools and FA strategies will be more beneficial in terms of motivation, confidence, and depth of learning, which is specifically useful in the postgraduate professional development.

5. Teachers' Perceptions and AL

The key role in changing to the new assessment paradigms is played by the attitude and skills of teachers. According to LAL research, the lack of formal training in assessment has a role in influencing the practice of teachers (Weng & Shen, 2022). Recent research demonstrates the mixed comprehension levels of AL by Iranian EFL educators. According to a nationwide survey, writing instructors regard themselves as assessors of writing, yet do not feel confident in evaluation, indicating the lack of training (Weng & Shen, 2022). Weigle (2007) also stated that the majority of the graduate TESOL programs do not include an assessment course, which makes teachers unprepared (and in some cases, they do not plan any assessment at all). Conversely, AL is not fixed; Amirian (2025) conducted the experiment and reported that more experienced EFL teachers have a better understanding of FA than the inexperienced

ones. Trainers found that experience and interpersonal skills (such as rapport with students) are of great influence on their FA skills (Amirian, 2025). It indicates that professional instructors, due to on-the-job learning, are likely to become progressively skilled in providing feedback and employing FA.

The assessment preference of a teacher is largely driven by the teacher's beliefs. Borg (2003) envisioned teacher cognition to consist of beliefs and knowledge to drive classroom work. Studies conducted on the Iranian context revealed that the conception of assessment, in general, is not uniform (i.e., clean emphasis on improvement vs. stringent accountability) among EFL instructors of different ages; genders (Forozandehfar & Manoochehrzadeh, 2024). Interestingly, in their survey, older and female teachers were found to be more likely to use formative strategies compared to male teachers and young teachers. The indicated demographic effect has been duplicated elsewhere: research conducted in China and Turkey indicates that factors related to the educational system (exam orientation) and practical limitations (class size, the curriculum load) prevent the use of FA (Zeng & Huang, 2021; Izci et al., 2024). In Iran, there were no particular barriers in these surveys, although culture-related prioritizing of exam scores may influence the culture in beliefs about summative measures.

6. Post-Pandemic Trends and Future Directions

Today, as institutions embrace a new normal, assessment practices keep changing. Blended teaching, where the real and online results are combined, is being retained in many universities. To that extent, hybrid assessment models are coming along. An

example here is that language programs might incorporate online quizzing as well as peer discussion boards as a component of final grades, as opposed to written tests. Vogt and Tzagari (2024) emphasized that emergency remote language evaluation appears to have provided numerous opportunities to language instructors interested in shifting the practices towards more FA and introducing hybrid types of learning and assessment. On a practical level, the courses could alternate between in-person presentation and virtual group activities, and therefore, they will offer different assessment scenarios. It is probable that authentic assessments (projects that resemble real-world tasks) will continue; the pandemic demonstrated that the assessment tasks, such as digital storytelling, e-portfolios, or service-learning projects, could be performed online and appreciated as a part of the final evaluation.

In the future, the use of AI-enhanced tools is becoming popular. Early evidence indicates that AI can provide a prompt, customized evaluation of written or spoken language (e.g., automatic grammar checking, a chatbot to help practice communication). To give an example, language testing professionals foresee few doubts about the use of AI-driven testing, as it may minimize bias and improve engagement. According to Biju et al. (2024), AI-aided writing tests reduced language anxiety and to a greater degree, motivation as compared to paper tests. Nonetheless, the evidence is still growing, and there is something to worry about (reliance on machine feedback or academic plagiarism (e.g., students using ChatGPT misused)). However, designers of the syllabus in English learning are discussing the way to incorporate AI into their syllabus (e.g., provide students

with AI-generated models as a learning activity or use it to differentiate tasks).

The important concern is whether such blended/formative practices are sustainable and scalable. A lot of positive developments have taken place during the crisis; however, will they withstand? Our research indicate hope intermingled with caution. Teaching professionals have built support systems on the one hand (Chung & Choi, 2021) and innovations themselves (e.g., online portfolios) have proven to be pedagogically valid. Conversely, unending problems associated with technology and the absence of systemic adjustment may cause a return to the focus on summative norms. Vogt and Tzagari (2024) stress that it is necessary to learn about the effects that COVID will have on language assessment on the periods of the pandemic and directly after it. They observe that it is important to note the lived emergency remote language assessment experience of teachers in establishing what practices you can still retain. Besides, teachers can use different assessment tools (portfolios, peer review) even in hybrid classes in case their LAL and digital skills have been improved during COVID. On the contrary, in case the institutions lack enabling practices for blended learning (e.g., stable LMS infrastructure, training in technologies), the post-pandemic tendency may shift towards the traditional forms of examination.

A lot of researchers support the Idea of going toward authentic and student-centered assessment on a permanent basis. As the pandemic demonstrated, student learning is positively influenced by the fact that the teacher has agency to design process-oriented tasks (521, 2021). After the pandemic, we will likely see further studies and policy attention as to the question of how this can be maintained. As an example, assigning

professional development points to the process of developing innovative assessments or curricula redesign in order to focus on critical thinking activities could be considered as a course of action. Interest in internationalizing AL occur globally as well; TESOL and university agencies are interested in training teachers with modules about e-assessment and so on. The Iranian situation may follow hybrid models with asynchronous tools such as e-portfolios and synchronous activities such as online presentations working with distance students. To conclude, post-pandemic trends predict the gradual shift towards an integrated, learner-centred assessment that would allow harnessing the technology (including AI) without compromising rigorousness and equity by abandoning pure summative exams. The scaling of these innovations countrywide will be determined by the persistence of research, policy and investment in infrastructure.

Although a rich COVID-related literature is available, there are still gaps that appear, in particular, in the context of TEFL in Iran. As an example, one of the few perspectives on experiences of Iranian EFL postgraduates is offered by Esfandiari et al. (2022), who discovered the feelings of unmet academic writing needs and inconsistent evaluation criteria among students. This leads to a relatively unexplored field, that is, how to evaluate high-level EFL proficiencies (academic writing, research). Since the TEFL programs are central to the training and research of teachers, it is essential to attain a sense of their evaluation context when planning curriculum and accreditation.

Therefore, it is time to study emergent summative/formative approaches with Iranian EFL teachers. It is able to fill a gap as it records what is done, thus associating it with

international trends and knowledge about training requirements. Specifically, the successes of the instructors can serve as guidance on the policy: depending on what instructional problems or successes teachers describe in the light of hybrid exams (e.g., whether they succeeded using workshops on LMS design), administrators could inform the support premises (e.g., perform workshops on LMS design in teaching). Considering that Popham (2009) cautions that inadequate LAL can result in inadequately structured language measurements and students themselves require more relevant measurements, there is a need to conduct research in order to tie the practices to intended pedagogical objectives. To conclude, the related research is justified due to the lack of exploration of postgraduate TEFL assessment in Iran, but its importance is crucial to the enhancement of academic and professional results in the digital age.

In particular, the current research strives to: compare pre-pandemic assessment policies and the levels at which they were based on traditional summative assessment options; discuss the use of assessment means and strategies before and during the pandemic, and the way they have changed; and query the EFL instructors on the advantages and inconveniences of using remote assessment in the TEFL world. Due to an emphasis on FA versus SA practices and by including the thoughts of teachers, the study becomes part of the continuously emerging global discussion on educational transformation, teacher growth, and digital AL in the post-pandemic period. Therefore, this research study proposes the following research questions in order to guide the investigation:

1. How did Iranian teachers assess EFL students prior to the COVID-19

- pandemic, and what strategies were commonly employed?
2. What assessment strategies were adopted during the COVID-19 pandemic for EFL students, and how did they differ from pre-pandemic methods?
 3. What are EFL teachers' perceptions of remote or e-assessment approaches introduced in the post-COVID-19 pandemic, particularly regarding challenges and potential benefits?

Method

The methodology adopted in this study was qualitative as it sought to understand the perceptions and practices of Iranian EFL teachers as participants in pre-pandemic, pandemic, and post-pandemic assessment strategies on TEFL learners. Since the proposed study aims at examining the lived experience of the participants, the most suitable design was identified to be phenomenological. Semi-structured face-to-face interviews have been used as a data collection instrument, which has made it possible to attain depth and breadth in the responses of the participants. Purposive sampling technique was retained to make sure that the respondents have pertinent experience in digital and traditional assessment situations. Analytical methodology was inductive-thematic, whereas the rigor was addressed through member checking, peer debriefing, and data triangulation. The research design, the selection of the participants, data collection instruments and methods, and analytical strategies are described in more detail in the following subsections.

1. Participants

In this study, twenty-five male and female EFL teachers in different institutions in Iran

took part in the research. The sample was chosen using purposive sampling as it is a strategy widely used in qualitative studies in order to select information-rich cases that closely match the purpose of study (Jones et al., 2006). The approach allowed incorporating educators possessing definite qualities that were related to research, i.e., (1) engaged in teaching EFL area, (2) had previous experience in utilizing digital tools in teaching, and (3) had an experience of at least five years of teaching experience. There was some attempt to produce gender, geographic, and digital literacy variation. Data saturation was used to determine sample size, which was a situation when no new ideas about themes were found during successive interviews or questionnaires (Saunders et al., 2018). It was saturated following 25 participants as it is recommended in phenomenological studies. The age of participants varied between 23 and 57 years. In order to preserve anonymity, every respondent received a numerical code. The geographical representation was also taken care of in transgressing the interviewees belonging to numerous institutes and universities in various cities of Iran. Both public and private ones had been involved to diversify the data with institutionality.

2. Instruments

The present study collected qualitative data with the help of face-to-face, semi-structured interviews, i.e., a qualitative data collection instrument that was adjusted and finalized in accordance with the previous studies conducted by Farhady and Tavassoly (2021): in particular, based on their formulated checklists. This conformity provided uniformity of data sources and increased reliability and comparability of the participants' responses. Interviews, depending on the needs of the research, can be

structured, unstructured, or semi-structured. The given research used a semi-structured interview design, which represents the compromise between consistency and flexibility. This type of format enabled the interviewer to use a pre-designed list of open- and closed-ended questions, but it also enabled an interviewer to define ambiguities and ask more questions relying on the answers of participants. Unscripted questions, in turn, will provide variety and give a chance to be spontaneous, as in the cases when it is not possible to predict what possible answers can be given (Dornyei and Taguchi, 2009).

The semi-structured interview protocol in this study was elaborated with regard to the research questions and based on previously read or known literature. Two experts in educational technology and applied linguistics examined the protocol independently, and some of the items on this protocol were reworded to eliminate ambiguity and enhance clarity. A pilot study was conducted with two EFL teachers, which resulted in some slight changes in the questions. The sample checklist of the finalized interview was comprised of four parts with first collecting demography information (e.g., age, gender, institution, and experience teaching at various levels); the second brief summary of the research objectives; the third containing open-ended questions concerning the major themes of perception, practice, and challenge regarding pre-, during, and post-COVID 19 assessments; and the fourth a place to leave any other comment.

3. Data Collection Procedures

The researcher collected the data within five months, from December 2024 to April 2025. The procedure was initiated in Iran. Institutional diversity was taken into

consideration by selecting the state-run and the private institutions. But in each interview, they were given a precise explanation of the aims of the study so that they could participate in the study without arbitrariness being avoided. All the participants gave their informed consent, and confidentiality and anonymity were assured. Approval of a university institutional review board was taken on ethical grounds. Where participants like those at Shiraz University refused audio recording, manual recording was done in real time through a structured interview checklist.

Data collection activity was carried out in alignment with the inductive logic of qualitative inquiry (Creswell et al., 2003), as answers to open questions generated rich, descriptive data gathered during interviews and observations. The data are then coded and grouped into themes, which enables researchers to characterize bigger trends and come up with grounded theoretical implications. The iterative pattern consisting of data collection, thematic classification, and theory-building also helped in holding the results close to the lived perceptions of the participants and illustrative of the complexity of reality faced during pre-pandemic and post-pandemic TEFL assessment. Remarkably, some aspects of grounded theory were introduced, where the goal was to come up with a theory based on information provided by the participants (Long et al., 1993).

All the interviews started with a short introduction on the purpose of the study and why the research was being completed to make sure that participants understood the topic of the research and were interested in it. In order to establish a connection and open dialogue, the researcher revealed to the participants his interest in the topic. Although interviews were structured based on preset

interview questions, a certain level of flexibility was observed in order to give participants a chance to speak with as much freedom as possible, in order to gather more data and a greater understanding of their experience regarding the assessment strategies.

With the aim of reducing bias and having as naturalistic data as possible, the researcher attempted first to audio-tape the interviews for a more precise transcription and analysis. Nevertheless, some of the participants refused to be videoed for privacy reasons. Under these circumstances, the researcher recorded the responses of participants during the research manually using an interview checklist prepared beforehand. In other cases, the respondents have chosen to fill in their answers in the interview forms to conserve time and as a move to be comfortable. These manipulations, though needed, were well controlled so as not to compromise the integrity and depth of the data that was gathered. Besides, the researcher was able to reflect deeply on his positionality to curb bias. The researcher was a TEFL scholar who had experience in applying various assessment tools in Iranian Universities and language schools. In order to reduce confirmation bias, he kept these reflexive journals when collecting data, where the researcher recorded assumptions (e.g., an initial assumption that the assessment tools would significantly decrease the amount of work). The highest degree of analytical rigor was also achieved through peer debriefing of two external qualitative researchers (Lincoln & Guba, 1985).

4. Data Analysis

The present study employed a qualitative research design, with data derived from

participants' responses to semi-structured interview questions as outlined in the instrumentation section. To address the research questions, thematic analysis was adopted as the principal method of data interpretation. To ensure the credibility and rigor of the thematic analysis, member checking and low-inference interpretation were systematically implemented, following established qualitative research standards (Lincoln & Guba, 1985; Braun & Clarke, 2006). Member checking involved sharing synthesized thematic summaries with six randomly selected participants to verify the accuracy of interpretations and alignment with their intended meanings. This process occurred approximately two weeks after initial data collection, with summaries sent via email in Persian to ensure accessibility. Participants were asked to confirm or clarify the themes, leading to refinements. This iterative feedback loop enhanced the authenticity of the findings, as advocated by Creswell and Creswell (2005).

Low-inference interpretation was prioritized to minimize researcher bias and anchor findings in participants' own words. During coding, *in vivo* labels were used for 70% of initial codes to preserve participants' language, reducing interpretive distortion (Braun & Clarke, 2006). Verbatim quotes were included extensively in the results to ground themes in raw data, with translations into English reviewed by a bilingual applied linguistics expert for fidelity. Reflexive journaling was maintained throughout the analysis, documenting assumptions and their evolution, as recommended by Guba and Lincoln (1994). Weekly peer debriefing sessions with two external qualitative researchers further mitigated bias, with 30% of transcripts independently coded to achieve 88% intercoder agreement ($Kappa = 0.85$).

Discrepancies were resolved through consensus discussions, ensuring analytical rigor. These strategies collectively ensured that the thematic findings were credible, transparent, and closely tied to participants' lived experiences.

Results

This section presents the findings derived from a thematic analysis of semi-structured interviews, organized around the study's research questions and reflecting Iranian EFL teachers' perceptions, practices, challenges, and contextual factors in evaluating assessment strategies for postgraduate TEFL learners pre- and post-COVID-19. Data were transcribed in Persian, iteratively reviewed, and semantically coded to identify themes and sub-themes, refined through constant comparison to capture salient patterns. Key themes and quotes were translated into English, with participants anonymized (e.g., P3, Q12). Additional quotes have been integrated to vividly illustrate teachers' perspectives, ensuring findings are deeply rooted in their voices.

1. Results for the First Research Question

To explore how Iranian EFL teachers assessed students before the onset of the COVID-19 pandemic and to identify the strategies most commonly used, participants were asked to reflect on their practices across different instructional levels. Analysis revealed two major themes: (1) Predominance of SA in Undergraduate Courses, and (2) Greater Pedagogical Flexibility and Research Orientation in Postgraduate Contexts. These themes were validated through member checking with six participants, all of whom confirmed that the categorizations accurately reflected their experiences. A low-inference

interpretive approach ensured that participants' intentions and phrasing were faithfully represented.

1.1. Theme 1: Predominance of SA in Undergraduate Courses

Preference for Final Exams Due to Class Size and Time Constraints: Many instructors reported that large class sizes necessitated the use of traditional final exams as the primary method of assessment. These exams were viewed as practical, time-efficient, and fair. One participant noted, *"It was hard and time-consuming for me to ask all students to have presentations or conferences during the class."* (Q2). Another added, *"The reason why I chose to assess my students in SA form is that, when it is impossible to ask all students to have a presentation or term project and to observe all students, it is somehow unfair for those students who cannot participate in such activities."* (P5). This preference reflects a pragmatic compromise between pedagogical ideals and classroom realities.

Student Motivation Shaped Assessment Practices: Participants frequently noted that students were predominantly grade-focused and showed limited interest in formative or project-based assignments. This influenced teachers to prioritize summative formats. A teacher elaborated, *"For most of my students, their final scores were important, and their priority was to pass the course. When I gave them projects to do, they did not do them carefully. So, I decided to just have a final exam to assess them."* (P7). This sentiment was shared by several instructors who observed minimal student engagement when assessments lacked direct influence on final grades.

Concerns About Objectivity and Feasibility in Evaluating Class Participation: Instructors

expressed hesitation to use class activities or presentations as core assessment components due to challenges in maintaining evaluation consistency and objectivity in large classes. One noted, *“Evaluating students’ presentations, term projects, and performance during the class was not only time-consuming but also difficult for me because I was not sure about the objectivity of my evaluation.”* (Q9). This concern contributed to the dominance of summative testing, which was viewed as more standardized and equitable under institutional guidelines.

Limited but Emerging Use of Formative and Peer Assessment: While not widespread, some instructors experimented with peer feedback and in-class formative strategies, particularly in efforts to boost engagement. A teacher explained, *“I usually ask my students to give presentations and conferences, and evaluate other students’ presentations. At first, they didn’t find this method useful, but after two or three sessions, they adapted... They told me that they learned more from searching the topic on the internet and also from their classmates.”* (P4). This quote highlights the potential of formative strategies to support active learning, even in environments where such practices were not institutionally emphasized.

1.2. Theme 2: Greater Pedagogical Flexibility and Research Orientation in Postgraduate Contexts

Smaller Class Sizes Enabled Individualized Assessment: Unlike BA courses, postgraduate (MA and PhD) and private institute classes were typically smaller, affording instructors more flexibility to assign research projects, presentations, and other formative activities. One elaborated, *“Despite the BA courses, the number of students in MA and Ph.D. courses*

was lower, and assigning projects for each student was easier for me. Also... observing, evaluating, and giving feedback to students separately was easier and more effective.” (Q6). This allowed teachers to design assessments that aligned more closely with students’ academic development.

Emphasis on Research Skills Development: Teachers consistently integrated research projects into postgraduate coursework to prepare students for thesis and dissertation writing. Nearly half of the participants described, *“I always asked my postgraduate students to do research and give them projects to work on to prepare them for doing their thesis and dissertation.”* (Q3). P14 explained, *“A term project is an activity that helps students to increase their ability to conduct larger research. So, I always gave my students... different projects during their courses to help them do research easily.”* (Q1). This reflects a broader pedagogical goal of cultivating scholarly autonomy and critical inquiry at advanced academic levels.

Balanced Assessment Practices to Reduce Anxiety: Some instructors used a combination of summative and formative methods to alleviate exam stress and enhance learning outcomes. Approximately 60% of participants emphasized, *“In the courses where I assessed students by a combination of SA and FA, students were less anxious in the final exam. This is a positive point regarding using FA for assessing students.”* (Q8). These blended approaches appeared to foster a more supportive and engaging academic environment, particularly among motivated postgraduate learners.

Engagement through Integrated Course Activities: Teachers reported incorporating student-led presentations and collaborative learning tasks to enhance classroom

participation and comprehension. *“I usually had the main source for my exams, and another source that asked students to present it to the class. By doing so, my students participated in class activities, and in my opinion, their learning increased.”* (P6). Such methods were regarded not only as assessment tools but also as instructional strategies that promoted peer learning and independent study.

Overall, prior to the COVID-19 pandemic, assessment strategies diverged significantly between undergraduate and postgraduate TEFL programs. At the BA level, institutional norms, class size, and logistical constraints led to a reliance on traditional SAs. In contrast, postgraduate instructors leveraged smaller class sizes and academic expectations to employ more research-oriented, formative, and individualized assessment strategies. These themes were confirmed through member checking, and participants agreed that the thematic distinctions accurately captured the practical and pedagogical rationale behind their assessment decisions.

2. Results for the Second Research Question

To examine how assessment strategies changed during the COVID-19 pandemic and how they diverged from pre-pandemic practices, participants were asked to compare their pre- and mid-pandemic experiences. Thematic analysis revealed three core patterns: (1) Initial Reliance on Traditional SA, (2) Adoption of Blended Assessment Models, and (3) Assessment Design Adjustments Due to Contextual Constraints. These themes were confirmed through member checking with six participants who reviewed the thematic summaries and agreed that they accurately captured their experiences. The findings are presented using

a low-inference interpretive lens, preserving the integrity of participants' original meanings.

2.1. Theme 1: Initial Reliance on Traditional SA

Default Use of Final Exams Amid Technological Unfamiliarity: At the onset of the pandemic, many instructors defaulted to SAs, primarily final exams, due to limited familiarity with online tools and the perceived infeasibility of FA methods in asynchronous or offline contexts. More than two-thirds of respondents indicated, *“In the first months of the COVID-19 pandemic and lockdowns, we were unfamiliar with new technologies, and some of our classes were held offline. In offline classes, it is hard to assess students by FA methods, so the best method, in my opinion, was taking a final exam at the end of the course”* (Q2). These early approaches reflected a continuation of pre-pandemic assessment norms, adapted only slightly to accommodate the remote format.

Continued Use of Summative Exams in Large Classes: Instructors teaching large BA courses found it more feasible to rely on SA due to scale and efficiency. However, this preference sometimes extended to MA courses taught by the same instructors. Nearly 40% of teachers expressed, *“The reason why I chose SA was that the population of BA courses was more than postgraduate courses, and assessing them was more difficult.”* (P4). Such comments reveal the logistical inertia that influenced assessment design in the early transition period, even at the postgraduate level.

2.2. Theme 2: Adoption of Blended Assessment Models

Shift Toward Combining FA and SA for Fairness and Engagement: As instructors

became more comfortable with digital tools, many adopted hybrid models that incorporated both FA and SA. This shift was motivated by pedagogical as well as ethical concerns, chiefly the need to reduce stress and ensure fairness amid unreliable exam platforms. Half (n = 12) noted, *“In the first semester of virtual education, I designed only a final exam for my students, but when the exam started in the LMS, some students faced problems and could not submit their answers. After that, I changed my method and used a combination of FA and SA to make my assessments fair and reduce the students’ stress”* (P9). Blended approaches also helped offset the risk of grade inflation caused by uncontrolled online exam environments. Most (60%, n = 15) cited, *“During the pandemic period, the assessment method changed from face-to-face to e-assessment. In final exams, as SA methods, students could cheat during the exam, which affected the validity and reliability of the scores, so I decided to use both FA and SA to decrease the effect of the final exam on the final score of students”* (Q5).

Use of FA to Increase Participation and Accountability: In the absence of physical monitoring, instructors used participation-based FAs to incentivize student engagement in virtual classrooms. A quarter pointed, *“Controlling students in the virtual systems was difficult. I couldn’t be sure that students were present during the online session. So, I chose FA for assessing my students and told them that their class activities, participation in discussions, and doing projects were parts of their final score to force them to attend the class.”* (Q8). This strategy marked a pragmatic attempt to replicate classroom accountability in digital settings and was particularly effective in postgraduate contexts.

FA as Preparation for Research: Several instructors noted that FA, including term projects and literature reviews, helped students build foundational research skills necessary for their thesis work. One expressed, *“The most important reason for this decision was to help my students prepare for and get familiar with the process of doing research.”* (P3). Such long-term pedagogical goals gained prominence in postgraduate programs, where assessments were increasingly aligned with academic development rather than just course completion.

2.3. Theme 3: Assessment Design

Adjustments Due to Contextual Constraints

Project-Based Alternatives to Presentations: Due to technical issues, such as students' inability to share files or activate microphones, many instructors replaced oral presentations with individual term projects. More than two-thirds of respondents (n = 17) indicated, *“I usually asked my students to give presentations in the class and evaluate each other’s presentations. But in the online classes, some students were unfamiliar with the technology... So, I decided to give students term projects, and if students themselves wanted to have a presentation, I allowed them to present some parts in class.”* (Q6). Another said, *“Some of my classes were held offline, in which students could not have presentations, so I gave them term projects instead”* (P10). These substitutions, while reactive, enabled continuity in assessment while avoiding classroom disruption.

Technological Barriers Shaping Assessment Weighting: Frequent issues with online assessment platforms led many instructors to reduce the weighting of final exams and emphasize coursework and

continuous activities. Most teachers (85%, $n = 22$) valued, *“Because some students faced problems regarding the system of online assessment, or their answers were not submitted in the system, I decided to allocate more scores to the FA methods and students’ activities during the course.”* (Q1). Another added, *“I allocated fewer scores to the final exam during the pandemic because there was no control and observation during the online exam, and in my opinion, the students’ grades might lack validity”* (P6). This shift was driven more by contextual limitations than instructional philosophy, yet it encouraged wider adoption of continuous assessment strategies.

Simplified Test Formats for Practicality: To manage grading loads and accommodate system constraints, many instructors adapted test formats, often replacing open-ended items with machine-gradable multiple-choice questions. Nearly all emphasized, *“I designed questions in the form of multiple-choice questions because correcting open-ended questions was difficult for me. Also, it saved my time with the help of machine correction”* (Q7). However, this was selectively applied, with some postgraduate instructors avoiding multiple-choice formats due to heightened concerns over academic dishonesty. One explained, *“After the outbreak of the pandemic, I still used FA methods but decided not to use multiple-choice questions because the possibility of cheating for such questions was higher”* (P11).

Taken together, the pandemic catalyzed a shift in assessment strategies from traditional, exam-centered models toward more adaptive and balanced approaches, especially at the postgraduate level. Initial dependence on SAs gave way to blended practices that integrated formative elements to mitigate technological

and ethical challenges. Teachers tailored assessments not only to support student engagement but also to align with the digital constraints and pedagogical demands of remote education. These themes were validated through member-checking, with participants affirming the themes’ consistency with their classroom realities during the COVID-19 transition.

3. Results for the Third Research Question

To understand teachers’ perceptions of remote or e-assessment methods introduced after the pandemic, the study investigated their reflections on both the advantages and limitations of these approaches. Thematic analysis identified three overarching themes: (1) Perceived Benefits of E-Assessment Tools, (2) Concerns Regarding Assessment Validity and Academic Integrity, and (3) Technological and Infrastructural Barriers. Each theme includes subthemes derived through thematic coding and validated through member checking with six participants, who confirmed the clarity and accuracy of the thematic distinctions. A low-inference interpretive approach was adopted to preserve participants’ original meanings and minimize researcher bias.

3.1. Theme 1: Perceived Benefits of E-Assessment Tools

Increased Efficiency and Reduced Workload: Many participants highlighted how the integration of e-assessment tools, especially automated grading systems, streamlined the assessment process and saved time. These tools were particularly valued in large undergraduate classes. *“Using e-assessment systems’ different options, such as autocorrect and question randomization, helped me to save my time and focus more on*

assessing my students' class activities and giving them feedback" (P12). *"Designing multiple-choice exams in the online assessment systems made the teachers' job easier in marking students' answers"* (Q6). *"E-assessment after the pandemic period reduced many teachers' workload in marking short-answer and multiple-choice questions"* (P8). These efficiencies allowed some instructors to shift their attention from mechanical grading to more qualitative tasks like student feedback and formative evaluation.

Advancement in Digital Competency and Future Preparedness: Several instructors acknowledged that the pandemic catalyzed upskilling in digital assessment, which they believed would benefit both educators and students in the long term. *"Using technology in this period is useful because it prepares teachers and students for the future when technology will have a greater role in education"* (Q2). *"Knowing new technologies for education and online assessment is very useful"* (P5). This perception suggests that despite initial resistance, many teachers viewed e-assessment as a valuable opportunity for professional development and future-readiness.

Positive Impact on Learner Engagement through FA: Participants reported that the incorporation of FA, particularly in postgraduate contexts, enhances student engagement and contributes to a more learner-centered environment. *"The use of FA for postgraduate students makes them more activated and motivated in the class, which changes the class from teacher-centered to learner-centered"* (P9). This stands in contrast to earlier reliance on teacher-led methods, indicating a shift in classroom dynamics after the remote learning period.

3.2. Theme 2: Concerns Regarding Assessment Validity and Academic Integrity

Inability to Monitor Student Behavior During Exams: A dominant concern among instructors is the perceived lack of control in online environments, which undermines the reliability and fairness of assessments, particularly summative exams. *"I am not sure about the validity and reliability of the students' scores because I have no control over my students during the exam"* (Q3). *"In my opinion, the main challenge is the lack of control over students during the classes and exams. For those courses that use SA methods, such as final exams, the possibility of cheating and having high grades is so high"* (P6). *"Cheating is possible for students due to the lack of teachers' control over them."* (Q11). These sentiments reflect widespread anxiety about academic dishonesty in unproctored online environments, particularly when high-stakes SAs were used.

Reduced Trust in the Validity of Final Scores: Several instructors stated that, due to the challenges in verifying the authenticity of student performance, they reduced the weight of final exams in the overall grading schema. *"I allocated fewer scores to the final exam after the pandemic because there is no control and observation during the online exam, and in my opinion, the students' grades might lack validity"* (P4). *"Because of problems with the online exam systems or internet connection, some of my students cannot submit their answers in the system for final exams"* (Q7). This strategic reduction of summative weighting in favor of continuous assessment reflects a pragmatic approach to balancing fairness and feasibility.

3.3. Theme 3: Technological and Infrastructural Barriers

Technical Failures and Internet Instability: A recurrent theme is the fragility of the technological infrastructure, which frequently disrupts exam administration and class activities. Teachers expressed frustration at being unable to support students during these breakdowns. *“Many students and teachers face problems with the system or internet connection during the exam, which causes stress to the students and teachers”* (P10). *“Because of problems with the online exam systems or internet connection, some of my students cannot submit their answers in the system for final exams”* (Q1). *“Lack of appropriate technological devices and a qualified internet connection pose many problems to teachers and students after the classes and online exams”* (Q5). These challenges not only increase stress but also contribute to perceptions of inequity, especially for students from rural or lower-income backgrounds.

Constraints on Interaction and Flexibility: The limitations of online systems also restrict teachers’ ability to implement dynamic and interactive assessments. In cases where internet connectivity is unreliable, instructors have to abandon synchronous presentations and revert to more static project-based formats. *“Some of my students cannot activate their microphones or present their PowerPoint files, which are time-consuming for the class. So, I decided to give students term projects”* (P7). This shift is driven by necessity rather than pedagogy and further complicates the efforts to ensure engagement and fairness.

Teachers’ perceptions of e-assessment after the pandemic reflected a nuanced and multifaceted experience. While the use of digital tools is credited with increasing efficiency, reducing workload, and fostering

digital literacy, significant concerns persist regarding assessment validity, equity, and technological accessibility. The shift toward blended models that combined formative and summative strategies appears to have emerged as a compromise solution, balancing technological potential with the realities of limited oversight and infrastructure. These themes were confirmed through member-checking, where participants agreed that the results captured both the benefits and burdens of e-assessment during an unprecedented shift in educational delivery.

Discussion

This study examined the evolution of assessment strategies in postgraduate TEFL programs in Iran before, during, and after the COVID-19 pandemic, as perceived by EFL instructors. The findings indicate a distinct bifurcation in pre-pandemic practices: SA, such as final exams, predominated at the undergraduate level, while postgraduate instruction featured greater pedagogical flexibility and a research-oriented use of FA tools. However, the abrupt shift to online education during the pandemic initially intensified reliance on SA due to limited digital literacy and infrastructural constraints. Over time, instructors transitioned toward blended assessment models, integrating FA to enhance engagement, fairness, and alignment with evolving pedagogical goals.

Notably, instructors perceived FA as particularly conducive to postgraduate learners’ academic development, citing its role in cultivating research skills, critical thinking, and learner autonomy. At the same time, concerns about academic integrity, technical limitations, and inequities in access tempered enthusiasm for fully online SA, leading many educators to reduce the weight of final exams

and prioritize continuous assessment. The findings of this study provide a nuanced understanding of how Iranian university teachers assessed postgraduate TEFL learners before and during the COVID-19 pandemic, highlighting both continuity and transformation in assessment practices. The results align with and extend existing literature on language assessment, particularly in the areas of SA versus FA, AL, and the impact of digital tools in TEFL contexts.

The predominance of SA in undergraduate TEFL courses, driven by large class sizes and time constraints, resonates with prior studies in exam-driven EFL contexts. Vellanki et al. (2023) and Crusan et al. (2016) similarly noted that institutional and cultural pressures often prioritize high-stakes exams over formative approaches in Asian and Middle Eastern settings. The reliance on final exams, as reported by participants, reflects a pragmatic response to logistical challenges but also underscores a gap in AL, as described by Popham (2009) and Weng and Shen (2022). The limited use of FA at the undergraduate level, despite some experimentation with peer feedback, aligns with Remmi and Hashim's (2021) observation that teachers often underutilize FA's potential to enhance learning due to a lack of training or institutional support.

In contrast, the greater pedagogical flexibility in postgraduate contexts, characterized by smaller class sizes and a focus on research-oriented assessments, supports Weigle's (2007) argument that graduate TEFL programs require assessments that foster advanced skills like scholarly writing and critical inquiry. The emphasis on research projects and blended SA-FA approaches in MA and PhD courses aligns with Esfandiari et al.'s (2022) findings that postgraduate EFL

students benefit from formative feedback to meet academic writing demands. However, the variability in assessment practices, as noted by participants, suggests inconsistencies in evaluation criteria, echoing Sohrabi et al.'s (2022) critique of Iranian EFL assessment practices. This variability underscores the need for enhanced LAL among instructors, particularly in designing tasks that align with constructivist principles of active, contextualized learning (Vogt & Tsagari, 2024).

The shift from initial reliance on SA to the adoption of blended assessment models during the pandemic reflects a global trend toward more flexible and adaptive assessment practices in response to COVID-19 disruptions. The early dependence on final exams, driven by technological unfamiliarity, mirrors Momeni's (2022) findings that Iranian EFL teachers initially struggled with digital tools due to limited training and infrastructure. The subsequent adoption of blended FA-SA models, motivated by concerns about fairness and engagement, aligns with Chung and Choi's (2021) observation that the pandemic prompted instructors to explore process-oriented assessments like portfolios and peer reviews. This shift also supports Sutadji et al.'s (2021) argument that combining multiple assessment modes enhances authenticity in language evaluation.

The use of FA to increase participation and prepare students for research tasks in postgraduate programs highlights the adaptability of formative approaches in digital contexts, as noted by Al Kamli and Almalki (2024) in their discussion of LOA. The replacement of presentations with term projects due to technological constraints further illustrates the pragmatic adjustments described by Şenel and Şenel (2021), who

found that Turkish EFL instructors adopted project-based tasks to maintain assessment continuity during remote learning. However, the reliance on simplified test formats, such as multiple-choice questions, raises concerns about assessment validity, echoing Al-Bargi's (2022) findings on the tension between convenience and academic integrity in online writing assessments.

Teachers' perceptions of e-assessment tools reveal a complex interplay of benefits and challenges. The increased efficiency and reduced workload associated with automated grading systems align with Olasina's (2023) findings on the growing relevance of digital assessment tools post-COVID. The advancement in digital competency, as reported by participants, supports Ashrafian and Alimohammadi's (2024) call for targeted professional development in digital assessment practices, suggesting that the pandemic catalyzed upskilling. The positive impact of FA on learner engagement, particularly in postgraduate contexts, corroborates Ismail et al.'s (2022) findings that formative approaches reduce test anxiety and foster self-regulation skills.

However, concerns about assessment validity and academic integrity, particularly in unproctored online exams, are consistent with global literature. Momeni's (2022) study on Iranian EFL teachers' perceptions and Vellanki et al.'s (2023) exploration of academic integrity in remote assessments both highlight the difficulty of ensuring fairness in digital environments. The strategic reduction of final exam weighting in favor of continuous assessment reflects a pragmatic response to these challenges, aligning with Tian et al.'s (2021) conceptualization of COVID-19 as a "critical incident" that exposed limitations in traditional assessment frameworks.

Technological and infrastructural barriers, such as internet instability and platform glitches, further complicate e-assessment, resonating with Emelogu et al.'s (2022) findings on digital access inequities in ESL contexts. These barriers disproportionately affected students from rural or lower-income backgrounds, underscoring the need for equitable assessment solutions, as advocated by Vogt and Tsagari (2024).

The findings contribute to the theoretical understanding of AL and LOA in TEFL contexts. The variability in pre- and mid-pandemic practices highlights the role of teacher cognition, as conceptualized by Borg (2003), in shaping assessment choices. The shift toward blended and formative approaches during the pandemic supports the constructivist paradigm (Liu et al., 2021), which emphasizes assessment as a co-constructed, learning-oriented process. Practically, the study underscores the need for institutional support in the form of professional development, stable digital infrastructure, and clear assessment guidelines to sustain innovative practices post-pandemic. The positive reception of FA in postgraduate settings suggests that LOA frameworks could be integrated into TEFL curricula to enhance student engagement and academic development.

Despite its contributions, this study has limitations. The sample, while diverse, is limited to 38 participants, which may not fully represent the experiences of all Iranian EFL instructors. The reliance on self-reported data, while rich in qualitative depth, may be subject to recall bias, particularly for pre-pandemic practices. Additionally, the study focuses on postgraduate TEFL, leaving undergraduate assessment practices underexplored in depth. Future research could address these gaps by

incorporating larger samples, longitudinal designs, and mixed-methods approaches to triangulate teacher and student perspectives.

Conclusion

This study provides a comprehensive exploration of how Iranian university teachers assessed postgraduate TEFL learners before and during the COVID-19 pandemic, revealing a transition from summative-dominated practices to more adaptive, blended, and formative approaches. Pre-pandemic, undergraduate courses relied heavily on final exams due to logistical constraints, while postgraduate programs leveraged smaller class sizes to incorporate research-oriented and FAs. During the pandemic, initial reliance on summative exams gave way to blended FA-SA models, driven by technological constraints and concerns about fairness and engagement. Teachers' perceptions of e-assessment highlight both its potential (efficiency, digital upskilling, enhanced engagement) and its challenges (validity concerns, technological barriers), aligning with global trends in language assessment literature.

These findings contribute to the ongoing dialogue on assessment reform in TEFL, emphasizing the need for enhanced LAL, institutional support, and equitable digital infrastructure to sustain innovative practices. The shift toward LOA and blended models in postgraduate contexts offers a promising direction for fostering advanced language skills and scholarly autonomy. Future research should explore the long-term sustainability of these practices, investigate student perspectives, and examine the role of emerging technologies, such as AI-assisted assessment, in shaping the future of TEFL assessment. By addressing these areas,

educators and policymakers can better align assessment practices with the evolving demands of language education in the post-pandemic era.

References

- Al Kamli, H. M., & Almalki, M. S. (2024). Learning-oriented assessment (LOA) implementation in EFL tertiary contexts: Towards a more task-based learning (TBL) environment. *World*, 14(1), 1-8. <https://doi.org/10.5430/wjel.v14n1p1>
- Alahmadi, N., Alrahaili, M., & Alshraideh, D. (2019). The impact of the formative assessment in the speaking test on Saudi students' performance. *Arab World English Journal (AWEJ)*, 10(1), 259-270. <https://dx.doi.org/10.24093/awej/vol10no1>
- Al-Bargi, A. (2022). Exploring online writing assessment amid COVID-19: Challenges and opportunities from teachers' perspectives. *Arab World English Journal*, 8(2), 80-96.
- Ali, S.A., Baloch, M., Ahmed N., Ali A.A., & Iqbal A. (2020). The outbreak of Coronavirus Disease 2019 (COVID-19): An emerging global health threat. *Journal of Infection and Public Health*, 13(4), 644-646. <https://doi.org/10.1016/j.jiph.2020.02.033>
- Amirian, S. M. R. (2025). Understanding EFL teachers' formative assessment literacy: insights from a mixed-methods study. *Language Testing in Asia*, 15(1), 1-25. <https://doi.org/10.1186/s40468-025-00363-y>
- Ashrafian, M., & Alimohammadi, N. (2024). Exploring EFL Teachers' Perception of Professional Development Needs in Post-Covid Era: Does Gender Make any Difference?. *Journal of New Trends in English Language Learning (JNTELL)*, 3(5). <https://doi.org/10.71528/2024.1194612>
- Berry, V., Sheehan, S., & Munro, S. (2017). Exploring teachers' language assessment literacy: A social constructivist approach to understanding effective practices. In *ALTE (2017). Learning and Assessment: Making the Connections-Proceedings of the ALTE 6th International Conference, 3-5 May 2017*. (pp. 201-207).
- Biju, N., Abdelrasheed, N. S. G., Bakiyeva, K., Prasad, K. D. V., & Jember, B. (2024). Which one? AI-assisted language assessment or paper format: An exploration of the impacts on

- foreign language anxiety, learning attitudes, motivation, and writing performance. *Language Testing in Asia*, 14(1), 1-45. <https://doi.org/10.1186/s40468-024-00322-z>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21, 5-31. <https://doi.org/10.1007/s11092-008-9068-5>
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81-109. <https://doi.org/10.1017/S0261444803001903>
- Bozkurt, A., & Sharma, R. C. (2020). Emergency remote teaching in a time of global crisis due to the Coronavirus pandemic. *Asian Journal of Distance Education*, 15(1), 1-6.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., ... & Walker, K. (2020). Purposive sampling: complex or simple? Research case examples. *Journal of Research in Nursing*, 25(8), 652-661.
- Chung, S. J., & Choi, L. J. (2021). The development of sustainable assessment during the COVID-19 pandemic: The case of the English language program in South Korea. *Sustainability*, 13(8), 1-13. <https://doi.org/10.3390/su13084499>
- Creswell, J. W., & Creswell, J. D. (2005). Mixed methods research: Developments, debates, and dilemmas. *Research in Organizations: Foundations and Methods of Inquiry*, 2(1), 315-326.
- Creswell, J. W., Clark, V. L. P., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed. *Handbook of mixed methods in social & behavioral research*, 209, 209-240.
- Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing*, 28, 43-56. <https://doi.org/10.1016/j.asw.2016.03.001>
- Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research*. Routledge.
- Emelogu, N. U., Nwafor, C. K., Chigbu, G. U., Okoyeukwu, N. G., & Eze, K. O. (2022). Awareness, proficiency, and challenges in the use of emerging technologies by ESL university lecturers in the post-COVID-19 era. *Cogent Education*, 9(1), 84-96. <https://doi.org/10.1080/2331186X.2022.2084962>
- Esfandiari, R., Meihami, H. & Jahani, F. (2022). Exploring postgraduate EFL students' academic writing experiences and expectations: A dynamic narrative approach. *Teaching English as a Second Language Electronic Journal (TESL-EJ)*, 25(4), 1-24. <https://tesl-ej.org/pdf/ej100/a4.pdf>
- Farhady, H., Tavassoli, K. (2021). Correction to EFL teachers' perceptions and practices of their language assessment knowledge. *Lang Test Asia* 11(1), 1-31 <https://doi.org/10.1186/s40468-021-00146-1>
- Forozandehfar, L., & Manoochehrzadeh, M. (2024). EFL Teachers' Conception and Choice of Assessment Task: The Focus of Age, Gender, and Level of Teaching. *Journal of New Trends in English Language Learning (JNTELL)*, 3(5), 83-106. <https://doi.org/10.71528/2024.07291127593>
- Friesen, N. (2010). Education and the social Web: Connective learning and the commercial imperative. *First Monday*, 15(12), 25-34. <https://doi.org/10.5210/fm.v15i12.3149>
- Ghanbari, N., & Nowroozi, S. (2021). The practice of online assessment in an EFL context amidst the COVID-19 pandemic: Views from teachers. *Language Testing in Asia*, 11(1), 1-27. <https://doi.org/10.1186/s40468-021-00143-4>
- Glazer, N. (2014). Formative plus Summative Assessment in Large Undergraduate Courses: Why Both?. *International Journal of Teaching and Learning in Higher Education*, 26(2), 276-286.
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of qualitative research*, 2(163-194), 105.
- Hunutlu, Ş., Keleş, N., & Yildiz, M. (2023). An investigation of online language assessment practices of EFL instructors in higher education: A special focus on the post-COVID-19 profile of language assessment literacy in the Turkish context. In *9th International KTUDELL Conference: Language, Literature and Translation 25-26 September 2023, Trabzon, TÜRKİYE* (p. 242).
- Hunutlu, Ş., Keleş, N., & YILDIZ, M. An Investigation of Online Language Assessment Practices of EFL Instructors in Higher Education: A Special Focus on the Post-COVID-19 Profile of Language Assessment Literacy in the Turkish Context. In *9th*

- International KTUDEL Conference: Language, Literature and Translation 25-26 September 2023, Trabzon, TÜRKİYE* (p. 242).
- Isbell, D. R., Kremmel, B. & Kim, J. (2023). Remote proctoring in language testing: Implications for fairness and justice. *Language Assessment Quarterly*, 20(4- 5), 469-487. <https://doi.org/10.1080/15434303.2023.2288251>
- Ismail, S. M., Rahul, D. R., Patra, I., & Rezvani, E. (2022). Formative vs. summative assessment: impacts on academic motivation, attitude toward learning, test anxiety, and self-regulation skill. *Language Testing in Asia*, 12(1), 1-40. <https://doi.org/10.1186/s40468-022-00191-4>
- Izci, K., İltter, İ, & Izgar, G. (2024). Adapting the Teacher Formative Assessment Literacy Scale into Turkish: Validation and reliability study. *International Journal of Assessment Tools in Education*, 11(1), 67-87. <https://doi.org/10.21449/ijate.1343373>
- Jones, S. R. (2013). Negotiating the complexities of qualitative research in higher education. *Journal of Research on Technology in Education*, 4(3), 55-84. doi.org/10.4324/9780203123836
- Kigotho, W. (2020). Law students approach courts to challenge online exams. *University World News*. 1 July 2020. <https://www.universityworldnews.com/post.php?story=2020070108564996>
- Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Springer.
- Liu, M., Kitto, K., & Shum, S. B. (2021). Combining factor analysis with writing analytics for the formative assessment of written reflection. *Computers in Human Behavior*, 120(1), 106-133. <https://doi.org/10.1016/j.chb.2021.106733>
- Long, D. R., Strauss, A., & Corbin, J. (1993). Basics of qualitative research: Grounded theory procedures and techniques. *The Modern Language Journal*, 77(2), 221-235. <https://doi.org/10.2307/328955>
- Maaoui A., Tzagari D., and Dammak H. (2023). The Relationship between language educators' perceptions and assessment practices during the COVID-19 Crisis. *Languages* 8(1), 54 <https://doi.org/10.3390/languages8010054>
- Momeni, A. (2022). Online Assessment in Times of COVID-19 Lockdown: Iranian EFL Teachers' Perceptions. *International Journal of Language Testing*, 12(2), 1-24.
- Ngololo, E. (2021, March 3). *Impact of COVID-19 on the University of Namibia*. SADC Higher Education Webinar, UNESCO.
- Nicol, D. (2007). E-assessment by design: using multiple-choice tests to good effect. *Journal of Further and Higher Education*, 31(1), 53-64. <https://doi.org/10.1080/03098770601167922>
- Olasina, G. (2023). Using new assessment tools during and post-COVID-19. *Library Philosophy & Practice*, 12(2), 11-24.
- Pokhrel, S., & Chhetri, R. (2021). A literature review on the impact of the COVID-19 pandemic on teaching and learning. *Higher Education for the Future*, 8(1), 133-141. <https://doi.org/10.1177/2347631120983481>
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental?. *Theory into Practice*, 48(1), 4-11. <https://doi.org/10.1080/00405840802577536>
- Purpura, J.E., Davoodifard, M., & Voss, E. (2021). Conversion to remote proctoring of the community English language program online placement exam at Teachers College, Columbia University. *Language Assessment Quarterly*, 18(1), 42-50. <https://doi.org/10.1080/15434303.2020.1867145>
- Remmi, F., & Hashim, H. (2021). Primary school teachers' usage and perception of online formative assessment tools in language assessment. *International Journal of Academic Research in Progressive Education and Development*, 10(1), 290-303. <http://dx.doi.org/10.6007/IJARPED/v10-i1/8846>
- Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., ... & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity*, 52(1), 1893-1907.
- Şenel, S., & Şenel, H. C. (2021). Remote assessment in higher education during the COVID-19 pandemic. *International Journal of Assessment Tools in Education*, 8(2), 181-199. <https://doi.org/10.21449/ijate.820140>
- Sohrabi, Z., Ghanbari, N., & Abbasi, A. (2022). Iranian EFL teachers' perceptions of writing assessment literacy: a countrywide study. *Language Testing in Asia*, 12(1), 1-23. <https://doi.org/10.1186/s40468-022-00172-7>
- Sutadji, E., Susilo, H., Wibawa, A. P., Jabari, N. A. M., & Rohmad, S. N. (2021). Adaptation strategy of authentic assessment in online learning during the COVID-19 pandemic.

- Journal of Physics: Conference Series*, 1810(1), 1-8. <https://doi.org/10.1088/1742-6596/1810/1/012059>
- Tian, W., Louw, S., & Khan, M. K. (2021). COVID-19 as a critical incident: Reflection on language assessment literacy and the need for radical changes. *System*, 103, 26-42.
- Vellanki, S. S., Mond, S. & Khan, Z. K. (2023). Promoting academic integrity in remote/online assessment - EFL teachers' perspectives. *Teaching English as a Second Language Electronic Journal (TESL-EJ)*, 26(4), 1-21. <https://doi.org/10.55593/ej.26104a7>
- Vogt, K., & Tsagari, D. (2024). Insights from emergency remote language assessment for a post-pandemic world. *Studies in Language*, 13(2), 115-146.
- Weigle, S. C. (2007). Teaching writing teachers about assessment. *Journal of Second Language Writing*, 16(3), 194-209.
- Weng, F., & Shen, B. (2022). Language assessment literacy of teachers. *Frontiers in Psychology*, 13(1), 45-82. <https://doi.org/10.3389/fpsyg.2022.864582>
- Zeng, J., & Huang, L. (2021). Understanding formative assessment practice in the EFL exam-oriented context: An application of the theory of planned behavior. *Frontiers in Psychology*, 12, 774159. <https://doi.org/10.3389/fpsyg.2021.774159>