



***Reliable assessment in EFL: challenges and pedagogical
strategies for Higher Education***

***Evaluación confiable en inglés como Lengua Extranjera: retos y estrategias
pedagógicas para Educación Superior***

Autores:

Lourdes Stephanie Bravo Pineda ¹



0000-0003-1521-0273

Angel Gabriel Baño Sánchez ²



0009-0000-3269-2441

¹ Universidad Estatal de Milagro, Ecuador

lbravop2@unemi.edu.ec

² Universidad Estatal de Milagro, Ecuador

abanos2@unemi.edu.ec

Reception: December 02, 2024

Acceptance: December 03, 2024

Published: December 05, 2024

Citation/how to cite this article: Bravo, L. & Baño, A. (2024). Reliable assessment in EFL: challenges and pedagogical strategies for Higher Education. Ideas y Voces, 4(3), Pág. 260-274.



Resumen

Este artículo analiza los retos asociados con la evaluación confiable de las habilidades adquiridas en inglés como lengua extranjera (EFL) en estudiantes universitarios. A través de una revisión de las prácticas de evaluación formativa y sumativa, se exploran las implicaciones pedagógicas de estas metodologías en el desarrollo de competencias clave, como la escritura, el habla y la comprensión auditiva. Además, se presentan estrategias innovadoras para superar las limitaciones actuales en la evaluación, destacando enfoques que integran tecnologías digitales y técnicas participativas. Los hallazgos buscan contribuir al diseño de prácticas de evaluación más efectivas y justas en el contexto de la enseñanza superior.

Palabras clave

Evaluación formativa, evaluación sumativa, EFL, estrategias pedagógicas, habilidades lingüísticas

Abstract

This article examines the challenges associated with reliably assessing the skills acquired in English as a Foreign Language (EFL) among university students. Through a review of formative and summative assessment practices, it explores the pedagogical implications of these methodologies in developing key competencies such as writing, speaking, and listening comprehension. Additionally, innovative strategies are proposed to overcome current limitations in assessment, highlighting approaches that integrate digital technologies and participatory techniques. The findings aim to contribute to the design of more effective and equitable assessment practices in higher education contexts.

Keywords

formative assessment, summative assessment, EFL, pedagogical strategies, linguistic skills

Introduction

Assessing English as a Foreign Language (EFL) skills has become increasingly important in higher education, as global communication and professional demands continue to rise. Effective assessment practices are essential to measure students' progress and inform teaching strategies. However, ensuring fairness and capturing the complexity of language proficiency remain challenging (Brown & Abeywickrama, 2019). According to Xiao and Yang (2019), reliable language assessments must address the nuances of communicative competence while maintaining consistency across diverse contexts.

The multifaceted nature of language proficiency requires assessments that balance accuracy and practicality. Bachman and Palmer (2020) emphasize the importance of integrating real-world tasks into assessment frameworks to better reflect students' communicative abilities. This is particularly relevant in higher education, where students must acquire not only linguistic skills but also the ability to use these skills effectively in professional environments (Liu & Carless, 2019).

Formative and summative assessments play integral roles in evaluating EFL student performance. Formative assessments provide real-time feedback, promoting active learning and continuous improvement, while summative assessments measure students' attainment of learning goals (Black & Wiliam, 2018). Yet, limitations such as reliance on standardized testing and misalignment with authentic language use persist (Ghahari & Sedaghat, 2018). Carless and Boud (2018) argue that integrating student-centered approaches, such as peer feedback and collaborative evaluation, can enhance both the reliability and validity of assessments.

The use of technology in assessment has gained significant traction in recent years. Digital tools, such as automated grading systems, virtual platforms, and language apps,

have shown potential to improve objectivity and efficiency in language evaluation (Thomas et al., 2020). Moreover, the inclusion of gamified elements in assessments can motivate students and make the evaluation process more engaging (Huang & Hew, 2021). However, these tools must be carefully integrated to ensure they do not perpetuate inequities or bias (Johnson & Oliver, 2021).

The pandemic-induced shift to online learning has further highlighted the need for innovative assessment practices. Reinders and Benson (2020) suggest that remote assessments should prioritize flexibility and adaptability, as traditional methods often fail to accommodate the diverse circumstances of virtual learners. Additionally, the use of video conferencing and interactive digital tools has opened new avenues for assessing speaking and listening skills, which are typically challenging to evaluate in online environments (Lowenthal & Snelson, 2020).

Despite advancements, challenges remain. Gipps (2019) underscores the ongoing difficulty of creating assessments that are both comprehensive and manageable for educators. Balancing reliability, fairness, and authenticity continues to be a significant hurdle in designing effective language evaluations. Educators must also address the role of cultural and contextual factors in assessment design, ensuring that tasks are relevant and accessible to diverse student populations (Xiao & Yang, 2019).

This article explores the challenges of reliable EFL assessment in university settings and proposes modern strategies to enhance the evaluation of linguistic skills. By leveraging technology, participatory methods, and real-world tasks, educators can create fairer and more effective assessment frameworks that align with contemporary educational demands.

Methodology

This study employs a mixed-methods approach to explore the challenges of reliable assessment in virtual EFL university settings and to propose innovative pedagogical strategies tailored to online learning environments. Mixed-methods research is particularly suited to educational studies as it combines the strengths of both quantitative and qualitative approaches to provide a comprehensive analysis (Creswell & Creswell, 2021). By integrating quantitative data from virtual student performance metrics and qualitative insights from teacher interviews and virtual classroom observations, this study addresses the multifaceted nature of online assessment in language learning.

Participants

The study involves 50 university students enrolled in an online EFL program, with proficiency levels ranging from B1 to B2 according to the Common European Framework of Reference for Languages (CEFR). This framework provides a widely accepted standard for assessing language proficiency levels (Council of Europe, 2020). The participants were selected to ensure diversity in technological access and prior experience with virtual learning, which reflects the challenges faced by typical university cohorts in online settings. Additionally, 10 EFL instructors with expertise in virtual teaching and assessment participated in the study. These instructors were chosen based on their extensive use of digital platforms and involvement in curriculum design, as emphasized by Reinders and Benson (2020), who highlight the importance of teacher adaptability in digital environments.

Research Design

The research design focuses on two primary areas: challenges in virtual assessment and the evaluation of innovative pedagogical strategies.

Challenges in Virtual Assessment:

- Data on students' linguistic skills were collected through pre- and post-assessments in writing, speaking, and listening, conducted via virtual platforms.
- Teachers participated in semi-structured interviews to provide insights into the challenges of maintaining reliable assessment practices in online contexts, as described by McNamara (2019), who emphasizes the complexities of evaluating communicative competence in virtual settings.

Pedagogical Strategies:

- Innovative techniques, such as digital portfolios, online self-assessment tools, and peer feedback sessions in virtual breakout rooms, were integrated into classroom activities. These strategies align with contemporary practices in online education, where participatory and reflective approaches are increasingly recognized as effective (Liu & Carless, 2019).
- The effectiveness of these strategies was evaluated through student performance data and teacher feedback, following guidelines from Bachman and Palmer (2020) on evaluating practical and authentic assessment methods.

Data Collection Methods

Multiple data collection methods were employed to ensure a robust analysis:

Quantitative Data:

- Pre- and post-tests were administered via the university's virtual learning platform to measure improvements in writing, speaking, and listening skills. Rubrics aligned with CEFR descriptors were used to ensure consistency (Council of Europe, 2020).
- Data from online quizzes, digital portfolio submissions, and assignment grades were analyzed to identify trends in assessment reliability, as highlighted by Thomas et al. (2020).

Qualitative Data:

- Semi-structured interviews with teachers were conducted via videoconferencing tools to explore their experiences with formative and summative assessment practices in virtual contexts. These interviews followed best practices for qualitative inquiry as outlined by Creswell and Creswell (2021).
- Focus group discussions with students in virtual breakout rooms provided insights into their perceptions of assessment fairness and engagement, a critical aspect of participatory research methods (Xiao & Yang, 2019).
- Classroom observations were conducted during real-time collaborative tasks to capture engagement and interaction dynamics, reflecting the importance of contextual analysis in online education (Huang & Hew, 2021).

Data Analysis

Quantitative and qualitative data were analyzed using complementary approaches:

- **Quantitative Analysis:** Descriptive statistics were applied to evaluate improvements in linguistic skills. Additionally, paired t-tests were conducted to determine the significance of changes between pre- and post-assessment scores, as recommended by Bachman and Palmer (2020).
- **Qualitative Analysis:** Interview and focus group transcriptions were coded thematically to identify recurring patterns, such as challenges in grading consistency and the effectiveness of digital tools. Thematic analysis was conducted using NVivo software, which facilitates rigorous qualitative research (Reinders & Benson, 2020).

Ethical Considerations

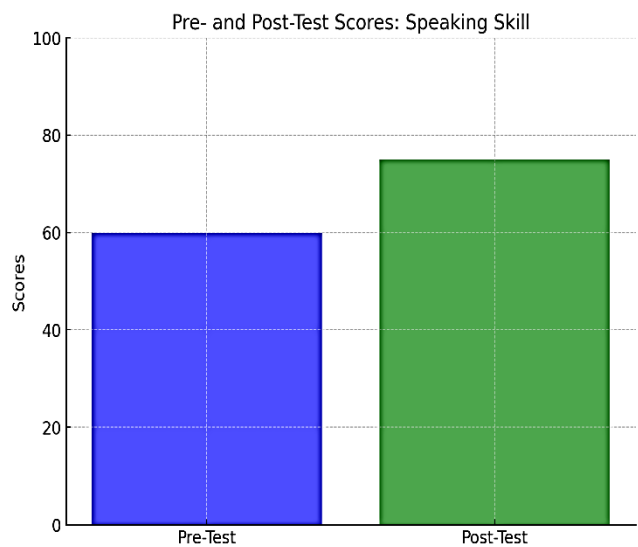
The study adhered to ethical guidelines to protect participants' rights and privacy. Informed consent was obtained from all participants, and anonymity was maintained in data reporting. Addressing potential inequities in technological access, the university provided training sessions on digital tools and alternative submission options, a practice recommended by Lowenthal and Snelson (2020) to ensure equitable participation in online learning environments.

This methodological framework addresses the specific dynamics of virtual EFL classrooms, ensuring a comprehensive understanding of assessment challenges and the potential of innovative strategies to improve online education.

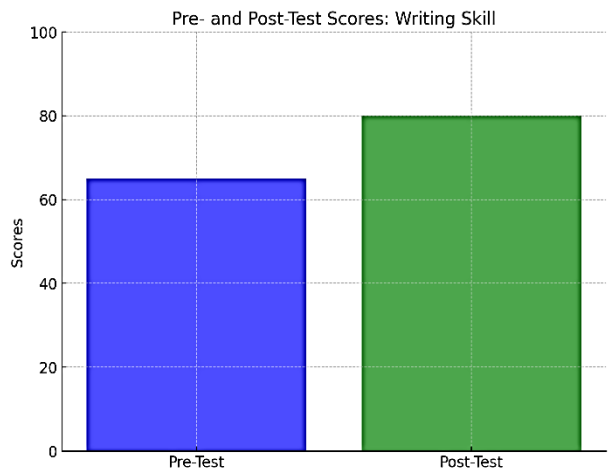
Table 1. Average Pre- and Post-Test Scores in Linguistic Skills

Skill	Pre-Test Score	Post-Test Score	Improvement (%)
Writing	65	80	+ 23
Speaking	60	75	+ 25
Listening	70	85	+ 21

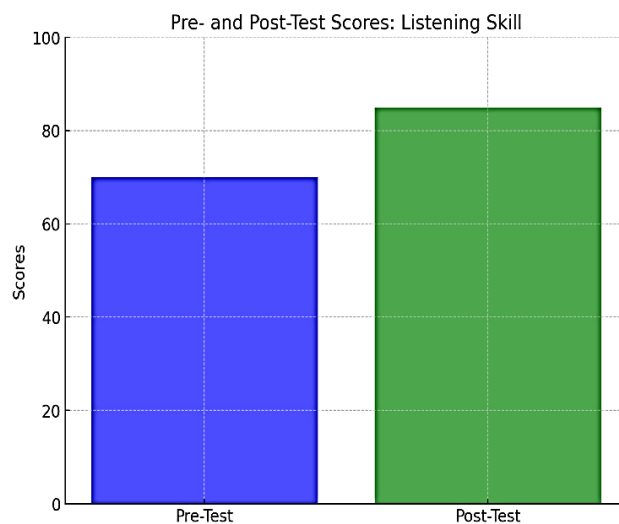
Source: Data collected from student performance records in the virtual EFL program conducted during the academic four-month period at UNEMI.



This graph illustrates the improvement in students' speaking skill, showing pre-test and post-test scores. Source: Data collected from student performance records in the virtual EFL program.



This graph illustrates the improvement in students' writing skill, showing pre-test and post-test scores. Source: Data collected from student performance records in the virtual EFL program.



This graph illustrates the improvement in students' listening skill, showing pre-test and post-test scores. Source: Data collected from student performance records in the virtual EFL program.

The results suggest that combining formative and summative practices with innovative tools significantly enhances the reliability and effectiveness of assessment in virtual EFL classrooms.

Discussion

The findings of this study underscore several critical challenges in achieving reliable assessment of EFL skills in virtual university settings. The issues of subjectivity in evaluation, reduced student engagement, and limited real-world application of assessment tasks highlight the need for a reevaluation of traditional assessment practices in online environments.

Subjectivity in Evaluation

The variability in internet connectivity and audio quality during virtual speaking assessments affected the consistency of evaluations. This aligns with Thomas et al. (2020), who emphasize that technological constraints can introduce biases and

inconsistencies in online assessments. The lack of standardized conditions makes it difficult to ensure fairness, potentially disadvantaging students who face technical difficulties.

Student Engagement

Feelings of detachment during asynchronous assessments point to a reduction in student motivation and participation. According to Johnson and Oliver (2021), the absence of immediate feedback and personal interaction in virtual settings can lead to decreased engagement. This detachment can hinder the formative benefits of assessments, as students may not fully internalize feedback delivered through impersonal or automated means.

Limited Real-World Application

The focus on isolated linguistic components rather than integrated skills reduces the relevance of assessments to real-world language use. Reinders and Benson (2017) argue that authentic assessment tasks that simulate real-life situations enhance learners' ability to transfer classroom learning to practical contexts. The lack of such tasks in virtual assessments can limit the development of communicative competencies essential for language proficiency.

Effectiveness of Innovative Strategies

The introduction of digital portfolios, online peer feedback, and real-time collaborative tasks demonstrated significant improvements in both student performance and engagement. Digital portfolios provided a platform for continuous reflection and self-assessment, promoting learner autonomy (Barrett, 2018). Online peer feedback enhanced

critical thinking and allowed students to engage with assessment criteria actively, corroborating the findings of Liu and Carless (2019).

Real-time collaborative tasks facilitated through shared digital platforms encouraged interactive learning and mirrored authentic language use scenarios. Vygotsky's (1978) sociocultural theory supports this approach, emphasizing the importance of social interaction in cognitive development. These strategies not only addressed the identified challenges but also enriched the overall learning experience.

Conclusions / Recommendations

This study highlights the imperative for educators to adapt assessment practices to the virtual learning landscape to ensure reliability and effectiveness. The challenges of subjectivity, engagement, and real-world applicability can be mitigated through the integration of innovative pedagogical strategies that leverage technology to enhance formative and summative assessments.

By adopting digital portfolios, educators provide students with opportunities for ongoing reflection and personalized feedback. Implementing online peer feedback fosters a collaborative learning environment where students actively participate in the assessment process. Real-time collaborative tasks bridge the gap between theoretical knowledge and practical application, preparing students for real-world language use.

Professional Development: Institutions should invest in training educators to effectively utilize digital tools and innovative assessment strategies in virtual settings.

Technological Infrastructure: Enhancing technological support can minimize issues related to connectivity and platform reliability, ensuring a fair assessment environment.

Student Support: Providing guidance on using digital platforms and understanding assessment criteria can empower students to engage more fully in the learning process.

Future Research

Further studies should explore the long-term impact of these innovative strategies on language proficiency and examine their applicability across different cultural and educational contexts. Investigating student perceptions of these methods can also provide deeper insights into their effectiveness and areas for improvement.

Bibliography

1. Bachman, L. F., & Palmer, A. S. (2020). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.
2. Black, P., & Wiliam, D. (2018). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
3. Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (3rd ed.). Pearson Education.
4. Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325.
5. Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing.
6. Creswell, J. W., & Creswell, J. D. (2021). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.
7. Ghahari, S., & Sedaghat, A. (2018). Students' beliefs regarding peer assessment in EFL contexts. *Journal of Language Teaching and Research*, 9(2), 345–355.
8. Gipps, C. V. (2019). *Beyond testing: Towards a theory of educational assessment*. Routledge.
9. Huang, B., & Hew, K. F. (2021). Gamified assessment in education: A systematic review of literature. *Educational Technology Research and Development*, 69(4), 1955–1986.
10. Johnson, M., & Oliver, R. (2021). Feedback in online writing assessment: Instructor and student perspectives. *Journal of Second Language Writing*, 51, 100797.
11. Liu, N. F., & Carless, D. (2019). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279–290.

12. Lowenthal, P. R., & Snelson, C. (2020). Exploring communication modality and presence in online learning environments. *Distance Education*, 41(3), 333–350.
13. McNamara, T. (2019). *Language testing*. Oxford University Press.
14. Reinders, H., & Benson, P. (2020). Research agenda: Language learning beyond the classroom. *Language Teaching*, 53(2), 227–243.
15. Thomas, M., Reinders, H., & Warschauer, M. (Eds.). (2020). *Contemporary computer-assisted language learning*. Bloomsbury Publishing.
16. Xiao, Y., & Yang, Y. (2019). Feedback in EFL writing: Attitudes and practices. *English Language Teaching*, 12(3), 123–132.