



Ensuring Fairness and Validity in EFL Oral Test Administration: A Socio-Cognitive Perspective from Vietnamese Universities

Nam Lam

Ho Chi Minh University of Banking, Vietnam

* Corresponding Author: **Nam Lam**

Article Info

ISSN (Online): 2582-7138

Impact Factor (RSIF): 7.98

Volume: 06

Issue: 06

November - December 2025

Received: 05-08-2025

Accepted: 07-10-2025

Published: 31-10-2025

Page No: 54-59

Abstract

This study investigates the administration of English oral assessment at tertiary institutions through the lens of Weir's (2005) socio-cognitive framework for test validation. It focuses on three contextual dimensions — physical conditions, procedural uniformity, and test security — to examine how institutional practices influence fairness and validity in speaking assessment. Using a convergent qualitative design, data were collected from 35 teachers, 352 students, six observed test sessions, supported by testing documents from three universities in South Vietnam. Findings indicate that while direct face-to-face formats foster communicative authenticity, inconsistency in examiner behavior and test-room setup may compromise reliability. The study highlights the need for systematic examiner training, improved physical conditions, and pre-test preparation to enhance the validity of EFL oral assessment in Vietnamese higher education.

DOI: <https://doi.org/10.54660/IJMRGE.2025.6.6.54-59>

Keywords: Oral Assessment, EFL Speaking Test, Test Administration, Socio-Cognitive Framework, Validity and Fairness

1. Introduction

Oral assessment plays a vital role in evaluating learners' communicative competence in English as a Foreign Language (EFL). As global standards in language testing increasingly emphasize authentic, performance-based evaluation, speaking tests have become an essential component of higher education programs. However, administering oral assessments reliably and fairly remains a complex challenge, influenced by test design, examiner behavior, and testing conditions (Taylor, 2011) ^[28]. Ensuring fairness and validity requires not only well-constructed tasks but also standardized administration procedures and transparent scoring practices (Weir, 2005) ^[31].

In Vietnam, English-speaking proficiency is central to university curricula, yet oral testing practices often vary across institutions. Research in Asian EFL contexts has similarly shown that insufficient examiner standardization and limited administrative resources lead to discrepancies in scoring and candidate performance (Isbell, Kremmel, & Kim, 2023; Liu, 2025) ^[17, 20]. Moreover, test-taker anxiety and environmental distractions can reduce the accuracy of performance evaluation, threatening the validity of results (O'Sullivan, 2012) ^[24].

This study investigates the administration of oral English tests across three tertiary institutions in Vietnam, focusing on how contextual and procedural factors influence the validity and fairness of speaking assessment. Drawing on classroom observations, questionnaires, and interviews with raters and candidates, the research explores shared practices and institutional differences in test conditions, examiner conduct, rating procedures, and perceptions of assessment reliability. Guided by Weir's (2005) ^[31] socio-cognitive framework, the study seeks to contribute practical insights into improving EFL oral test administration through enhanced examiner training, standardized procedures, and better-controlled testing environments.

Accordingly, this paper addresses the following research questions:

1. What contextual and procedural factors influence the validity and fairness of oral English test administration in Vietnamese tertiary institutions?
2. How do teachers and students perceive the effectiveness and challenges of current oral test administration practices?

2. A brief literature Reviews

In language assessment, the administration of oral tests plays a crucial role in ensuring validity and fairness. Within the socio-cognitive framework, context validity depends largely on three administrative factors—physical conditions, uniformity of procedures, and test security (Weir, 2005, p. 46) ^[31]. These elements shape the testing environment and influence how accurately a speaking test measures a candidate's communicative competence.

Physical conditions such as lighting, acoustics, room size, and seating arrangements can affect both candidate performance and examiner judgment. Research has shown that well-arranged and comfortable test rooms reduce anxiety and support more authentic interaction (Brown, 2003; O'Sullivan, 2012) ^[3, 24]. In contrast, poor acoustic quality, noise, or cramped seating may generate construct-irrelevant variance and reduce reliability.

Uniformity of test administration is essential for fairness and comparability across candidates. Variations in examiner behaviour, timing, or task delivery have been shown to produce scoring discrepancies (Galaczi & ffrench, 2011) ^[12]. Examiner training and clear procedural guidelines therefore contribute directly to reliability (Taylor, 2011) ^[28]. Studies on test format also reveal that the structure of interaction—one-on-one, paired, or group—influences assessment consistency. The one-on-one interview, while long-established (He & Young, 1998; Dandonoli & Henning, 1990) ^[7], may limit natural interaction (Li, 2019) ^[19]. The paired format aligns more closely with communicative language teaching (Fulcher, 2003) ^[10], though partner proficiency differences may influence outcomes (Nakatsuhara, 2004; Norton, 2005) ^[22, 23]. Group orals, despite efficiency in large-scale testing, raise concerns about fairness and scoring complexity, as interactional dominance or unequal participation can distort outcomes (He & Dai, 2006; Van Moere, 2006) ^[30].

Recent technological advancements further support standardization and fairness in oral assessments. Multimedia recording and computer-based testing (CBT) tools—including free software and web-apps—allow digital voice recordings that can be replayed for rater calibration, improving scoring consistency and feedback transparency (Early & Swanson, 2008; Pop & Dredetianu, 2013) ^[8, 25]. More recent work shows that AI-enabled mobile apps and digital peer-feedback platforms offer scalable and improved validity in oral assessment contexts (Liu, 2025; Alzubi, 2025; Weng, 2025; Stephenson, 2025) ^[20, 1, 32, 27]. These innovations reduce candidate anxiety, support self-assessment, increase uniformity of administration, and facilitate large-scale

implementation.

Test security, the third contextual factor in Weir's model, safeguards assessment integrity. This includes controlling access to test materials, verifying candidate identity, and maintaining confidentiality of performance data. Administrative roles such as ushers and coordinators help maintain order and prevent testing information leakage, contributing to consistent and ethical assessment practices.

Overall, the literature highlights the need for oral test administration to be conducted in a standardized, transparent, and secure manner to ensure validity. This research builds on these principles by examining how Vietnamese tertiary institutions apply them in English speaking assessments, drawing on data from surveys, interviews, and test-room observations.

3. Methodology

This section outlines the research setting and describes the methods used for data collection and analysis.

3.1. Research setting

Across the three universities in South Vietnam involved in the study, oral tests were administered on campus using existing facilities and teaching staff. All the institutions, namely University A, B, and C, employed direct assessment formats, with live interaction between candidates and examiners, and no use of audio or video recording. Raters typically decided scores immediately after each session without pre-test training or post-test moderation. These practices, while practical, posed risks to standardization and fairness in scoring (Kunnan, 2000; Gipps & Stobart, 2009) ^[18, 11].

The physical settings were ordinary classrooms rearranged slightly to accommodate examiners and candidates. Seating arrangements reflected different test formats: a one-on-one interview in University C, and paired-candidate sessions in Universities A and B. The paired configuration encouraged interaction and reflected communicative testing principles, but inconsistencies in examiner roles and placement sometimes led to candidate confusion and anxiety. Noise from waiting candidates was also noted as a distraction in some test rooms.

All institutions relied on in-house EFL teachers to serve as raters. While this practice was cost-effective and ensured staff familiarity with students, it limited inter-rater reliability due to the absence of calibration sessions. The lack of audio recording compromised transparency and limited opportunities for review or appeal when scores were contested.

3.2. Data collection and analysis

Data collection took place in three tertiary institutions that administered English oral summative assessments as part of their programs for English majors. Because of scheduling and institutional constraints, not all data were gathered simultaneously. Documentary materials on test formats and administrative procedures were obtained prior to the assessment period to support fieldwork preparation and

logistical coordination, such as candidate scheduling, voice recording, and the presence of co-investigators in multiple testing rooms.

In language assessment research, combining data from multiple sources provides a more comprehensive understanding of complex issues such as validity, rater effects, and assessment design (Moeller, 2016)^[21]. This study employed a convergent qualitative approach, collecting and analyzing different types of data separately but assigning them equal weight to strengthen the interpretation of results. Integrating multiple data strands enabled triangulation, allowing findings from one source to confirm or clarify those from another (Johnson & Christensen, 2012)^[16].

Both EFL teachers and students participated in the study. They not only provided information through separate questionnaire surveys but also took part as examiners and candidates in the oral tests. Their activities and behaviors were directly observed at the test administration sites to capture authentic testing practices. Table 1 summarizes the sample and instruments used. Data sources included questionnaire surveys (35 EFL teachers, 352 students), observations of six speaking test rooms, and recordings of 84 student speech samples. Institutional documents—such as course outlines, test banks, rubrics, and scoring sheets—were also reviewed to establish links between curriculum objectives and assessment practices.

Table 1: Sampling of participants and research instruments used for the study

Methods of data collection	Participants	N	Research instruments
Survey	+ EFL teachers	35	Questionnaire for teachers
	+ EFL students	352	Questionnaire for candidates
Interview	+ Individual EFL teachers	6	Protocol for teachers
	+ EFL students in 7 groups	27	Protocol for candidates
Observation	Speaking test rooms	6	Test-room observation scheme

Qualitative data were processed through content analysis of observation protocols, interview transcripts, and test materials to identify recurring patterns, meanings, and themes (Thomas, 2003)^[29]. The approach captured what occurred in the actual testing contexts and illuminated the relationship between test design, administration practices, and participants' experiences. This methodological design ensured that the study's conclusions were grounded in rich, contextually informed evidence from multiple perspectives.

4. Findings and discussion

4.1. Assessment procedures and examiner practices

Field notes from the observations revealed that the three institutions shared several similarities in oral test administration, but notable differences emerged in rating conditions and testing procedures. Table 2 summarizes their common practices and distinctive administrative features.

Table 2: Comparing test administration methods across institutions

Common features	All the three institutions		
- School-based using available facilities and human resources - Direct assessment (live scoring) - No audio-recording. Raters decided scores for individuals right after each speaking test session. - No rater training or meeting prior to the test for discussing what the assessment criteria were or how the raters would score and use the rating scale - No feedback on candidates' oral performance after testing			
Particular features	Uni. A	Uni. B	Uni. C
Interlocutor	-	✓	-
Usher	-	✓	-
Technology	✓	-	-
Air-conditioned	-	-	✓
Pre-informed test questions	-	-	✓
Timing	varied	varied	varied

All three institutions conducted oral assessments on campus using existing EFL classrooms and teaching staff. Test rooms were ordinary classrooms with only minor adjustments in seating to accommodate the number of examiners and candidates. While the physical conditions—lighting, space, and furniture—were generally adequate, seating layouts differed slightly due to format variations. University C adopted a one-on-one interview setting that enabled direct examiner–candidate interaction, whereas Universities A and B implemented a paired format involving two examiners and two candidates. The side-by-side arrangement of candidates encouraged balanced attention from the interlocutor and supported natural interaction (Cambridge English, 2016)^[4].

At University B, the distinction between the interlocutor and assessor roles reduced candidates' anxiety by clarifying examiner functions. However, at University A, seating both examiners side-by-side occasionally confused candidates about who would interact with them. University C employed a one-on-one interview format involving only two participants—an examiner and a candidate—throughout the entire test (Fan & Yan, 2020)^[9]. Overall, the observed arrangements enabled authentic communicative exchanges while reflecting institutional differences in test design and administration. Figure 1 illustrates the seating arrangements used by the institutions to facilitate candidates live oral performances through different patterns of interaction.

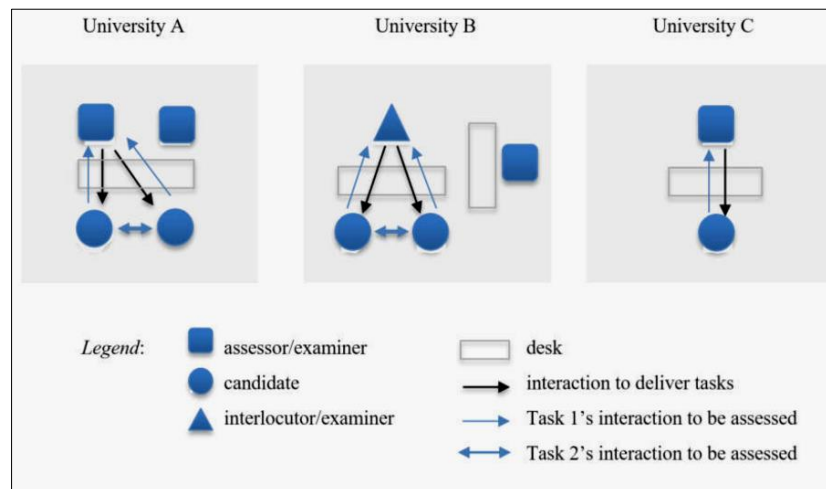


Fig 1: Seating arrangements for oral assessments across institutions

The oral tests followed a direct, two-way interaction model (Reed & Cohen, 2001) ^[26] involving either examiner-candidate or candidate-candidate exchanges. Examiners valued direct assessment for its authenticity and ability to elicit spontaneous language use. Candidates also preferred face-to-face formats over computer-based tests, finding them more motivating and communicatively engaging. However, live scoring demanded intense concentration from raters, increasing fatigue and the potential for subjective bias. Some raters had prior experience with semi-direct speaking tests from other institutions. One rater compared direct and semi-direct formats, expressing a clear preference for face-to-face scoring, in an interview:

Actually, marking from an audio-recording is not as good as direct marking where initially we can see their (candidates') facial expressions, then we can do the rating better than scoring an audio file. Sometimes it is difficult to completely understand what a candidate says in an audio file. However, we can guess the words a candidate is speaking when we look at his/her mouth shape. (U2.R1)

The absence of recording was one of the most significant weaknesses in the testing systems. Raters themselves recognized the benefits of recording oral performances for training, quality assurance, and pedagogical feedback. Despite awareness of these advantages, institutions avoided recording due to concerns about time, technical challenges, and data storage. As a result, speaking tests lacked documentary evidence comparable to written exams.

None of the institutions organized formal pre-test briefings to discuss scoring criteria or task delivery procedures. Instructions were typically delivered minutes before the test, leading to variability in timing, question phrasing, and interaction style. While some institutions included two raters—one interlocutor and one assessor—others relied on a single teacher-rater who both administered and scored the test. University B's dual-rating system, combining holistic and analytic judgements, was the most structured and contributed to score reliability (Green & Hawkey, 2012) ^[13]. The inclusion of an usher in one institution's setup was an effective administrative innovation. The usher coordinated candidate flow, minimized disruption, and allowed examiners to focus on their evaluation tasks. Another noteworthy variation was University C's practice of providing test questions weeks in advance to encourage student preparation, though this sometimes resulted in rehearsed or memorized responses that undermined communicative spontaneity.

4.2. Candidates' perceptions of oral test administration

Survey data with 352 students revealed generally positive perceptions of the oral test process but highlighted specific concerns. The majority valued teacher feedback highly ($M = 4.34$) and expressed a strong desire for post-test comments to guide future improvement (Table 3). This result aligned with nearly 90% of interview respondents agreed that examiner feedback helped them identify strengths and weaknesses.

Table 3: Means and standard deviations (SD) for candidates' perceptions of the oral test administration

Statements	N	Mean	SD
(1) I need teacher feedback on my speaking performance in the test so that I can improve my speaking skills.	352	4.34	.74
(2) I had sufficient time to prepare for this test.	352	3.68	.82
(3) I was clearly informed of the assessment criteria to prepare for the test.	352	3.62	1.02
(4) The test was well administered.	352	3.62	.86
(5) The atmosphere of the test room was stressful.	352	3.28	1.08
(6) I believe that computer-assisted speaking tests are more accurate than those by human raters.	352	2.42	.89
Valid N (listwise)	352		

Note: 1 = Strongly disagree 2 = Disagree 3 = Neither agree nor disagree 4 = Agree 5 = Strongly agree

Candidates felt adequately prepared for the tests ($M = 3.68$) but reported inconsistent awareness of assessment criteria. About 15% were unsure how scores were determined, reflecting the absence of standardized pre-test communication. Candidates were reasonable to be concerned about how different parts of the test were weighted, because

the score weight of each section or criterion shows how important it is compared with the others (Galaczi & ffrench, 2011, p. 127) ^[12]. A candidate shares her situation in an interview that she knew the test would be taken in pairs, but the assessment criteria were unclear. She was not sure how much grammar or sentence structure mattered, so she did not

know what to focus on when preparing (U2G1.S4). Some students reported stress and anxiety linked to the formal test atmosphere or examiner demeanor. While many candidates performed close to their true ability, others felt nervousness reduced their performance to 60–70% of potential. These findings align with research suggesting that test anxiety negatively affects cognitive output (Chuang, 2009) [6]. However, most of the students agreed that the test was well administered ($M=3.62$), which aligned with the test-room observation that no cheating or misidentification of candidates occurred. Candidates favored human-administered tests over computer-based ones ($M = 2.42$ for CBT preference). They appreciated the natural interaction and supportive gestures

from examiners, which encouraged active communication. However, inconsistent rater behavior and lack of feedback were noted as major limitations.

4.3. Raters' perceptions of oral test administration

Thirty-five EFL teacher-raters participated in the rater survey and follow-up interviews. They emphasized the importance of consistency and fairness in scoring ($M = 4.66$) and believed that students needed to master the language skills and content specified in the course objectives to perform well on the test ($M = 4.34$). As shown in Table 4, many raters acknowledged that oral tests effectively reflected the course objectives and helped assess integrated speaking skills.

Table 4: Means and standard deviations (SD) for raters' perceptions of the oral test administration

Statements	N	Mean	SD
(1) Ensuring consistency in rating and scoring are very important to me.	35	4.66	.59
(2) Students need to learn the language material and skills outlined in the course objectives to achieve good results for the test.	35	4.31	.58
(3) There was sufficient time for students to prepare for this test.	35	4.06	.76
(4) The atmosphere of the test room was formal.	35	3.54	.82
(5) I believed that computer-assisted speaking tests are more accurate than those by human raters.	35	2.29	.96
Valid N (listwise)	352		

Note: 1 = Strongly disagree 2 = Disagree 3 = Neither agree nor disagree 4 = Agree 5 = Strongly agree

Nevertheless, the time constraint for each speaking test session limited opportunities for providing feedback (U2.T3). Some raters believed feedback should be given later to avoid affecting students' emotions during exam periods (U2.T1). Others suggested written reports to guide future improvement. Raters also expressed skepticism toward computer-assisted oral testing, citing concerns about the loss of natural interaction, limited observation of body language, and technical reliability issues (Chapelle, 2008, p. 129) [5]. Interviews indicated that while most raters valued the authenticity of direct interaction, they recognized the risk of subjective bias and fatigue. The absence of calibration meetings meant that interpretations of rating scales varied across examiners, reducing scoring reliability. As Bøhn (2015) [2] noted, raters can diverge in how they prioritize assessment criteria even when using the same scales.

5. Discussion and implications

The findings suggest that oral test administration across Vietnamese universities shared practical strengths but lacked standardized protocols ensuring fairness and reliability. Direct interaction supported communicative validity and provided an authentic test context, yet the informal procedures undermined consistency. Institutional reliance on untrained in-house raters and absence of recordings made post-hoc moderation impossible and introduced subjective variation.

Students and teachers both emphasized the importance of transparency—particularly clear communication of assessment criteria and post-test feedback. Providing this information systematically could enhance learner confidence and align classroom teaching with assessment practices. The inclusion of supporting staff, such as ushers, improved efficiency and minimized distractions, offering a useful administrative model.

Although computer-assisted oral testing (CALT) presents long-term potential, most Vietnamese institutions are not yet equipped with adequate technical infrastructure or trained

personnel. As Chapelle (2008) [5] argues, sustaining technological testing systems requires continuous investment and staff development. Until then, human-mediated direct assessment remains the most viable approach, provided that institutions strengthen examiner training, implement double-rating procedures, and consider recording for accountability.

6. Conclusion and recommendation

This comparative investigation demonstrates that valid and fair oral test administration depends on coordinated collaboration between teaching and administrative staff, clear communication of assessment standards, and ongoing examiner development. The examined institutions achieved meaningful engagement between examiners and candidates through face-to-face testing, but issues of inconsistency, anxiety, and lack of feedback persist. Improvements in pre-test preparation, examiner calibration, and post-test reflection are essential for enhancing both validity and reliability of oral assessment in Vietnam's tertiary EFL programs.

Future research should explore the impact of examiner training and digital tools on scoring consistency, as well as longitudinal outcomes of feedback-informed oral assessment. Integrating technological innovations gradually, while preserving the communicative essence of speaking evaluation, will ensure that assessment practices continue to support language learning and educational quality assurance.

7. Thank-You Note

The author would like to express sincere appreciation to the EFL teachers and English-major student participants from the three universities in South Vietnam. Their valuable information and thoughtful insights formed the core of this study. The author also extends gratitude to previous researchers and scholars for their significant contributions to the field of language testing and assessment. Their work has provided essential theoretical and methodological foundations that guided the present investigation. It is hoped that the findings of this study will further enrich ongoing

discussions and research in language assessment.

8. References

1. Alzubi AAF. Unlocking the power of online assessments in EFL: teachers' and students' perceptions of oral evaluation. *SAGE Open*. 2025;15:1-12. doi:10.1177/21582440241311785.
2. Bøhn H. Assessing spoken EFL without a common rating scale: Norwegian EFL teachers' conceptions of construct. *SAGE Open*. 2015;5(4):2158244015621956.
3. Brown A. Interviewers, interlocutors, and testers. Cambridge: Cambridge University Press; 2003.
4. Cambridge English. First-level handbook: speaking test procedures. Cambridge: Cambridge University Press; 2016. Available from: <https://www.yle.tw/assets/file/picturebook/KET.pdf>.
5. Chapelle CA. Utilizing technology in language assessment. In: Shohamy E, Hornberger NH, editors. *Language testing and assessment*. 2nd ed. Vol. 7. Dordrecht: Springer; 2008. p. 23-134.
6. Chuang YY. A study of college EFL students' affective reactions and attitudes toward types of performance-based oral tests. *J Educ Res*. 2009;43(2):55-80.
7. Dandonoli P, Henning G. An investigation of the construct validity of the ACTFL oral proficiency interview. *Foreign Lang Ann*. 1990;23(1):11-22.
8. Early M, Swanson J. Multimedia recording for computer-based speaking tests: technological possibilities and validation issues. *Lang Test*. 2008;25(1):45-64.
9. Fan J, Yan X. Assessing speaking proficiency: a narrative review of speaking assessment research within the argument-based validation framework. *Front Psychol*. 2020;11:330.
10. Fulcher G. *Testing second language speaking*. Harlow: Pearson Education; 2003.
11. Gipps C, Stobart G. Fairness in assessment. In: *Educational assessment in the 21st century: connecting theory and practice*. Dordrecht: Springer; 2009. p. 105-18.
12. Galaczi E, French A. Context validity. In: Taylor L, editor. *Examining speaking: research and practice in assessing second language speaking*. Cambridge: Cambridge University Press; 2011. p. 112-70.
13. Green A, Hawkey R. Marking assessments: rating scales and rubrics. In: Coombe C, Davidson P, O'Sullivan B, Stoyneff S, editors. *The Cambridge guide to second language assessment*. Cambridge: Cambridge University Press; 2012. p. 299-306.
14. He AW, Young R. Language proficiency interviews: a discourse approach. In: Young R, He AW, editors. *Talking and testing*. Amsterdam: John Benjamins; 1998. p. 1-24.
15. He L, Dai Y. A study of group oral interaction in EFL testing. *Foreign Lang World*. 2006;5:387-92.
16. Johnson RB, Christensen L. *Educational research: quantitative, qualitative, and mixed approaches*. 4th ed. Thousand Oaks: SAGE; 2012.
17. Isbell DR, Kremmel B, Kim J. Remote proctoring in language testing: implications for fairness and justice. *Lang Assess Q*. 2023;20(4-5):469-87.
18. Kunnan AJ. Fairness and validation in language assessment: selected papers from the 19th Language Testing Research Colloquium, Orlando, Florida. Vol. 9. Cambridge: Cambridge University Press; 2000.
19. Li J. An evaluation of the IELTS speaking test: face-to-face interview format and its limitations. *Open Access Libr J*. 2019;6:1-17.
20. Liu XJ. Score accuracy, perceived validity, and oral peer feedback in AI-assisted speaking assessment tools. *Assess Educ Princ Policy Pract*. 2025; Advance online publication.
21. Moeller AJ. The confluence of language assessment and mixed methods. In: Moeller AJ, Creswell JW, Saville N, editors. *Second language assessment and mixed methods research*. Cambridge: Cambridge University Press; 2016.
22. Nakatsuhara F. The impact of pairing candidates of different proficiency levels in a paired speaking test [doctoral dissertation]. Colchester: University of Essex; 2004.
23. Norton J. The paired format in oral testing. *ELT J*. 2005;59(3):287-97.
24. O'Sullivan B. *Assessing speaking*. Cambridge: Cambridge University Press; 2012.
25. Pop IL, Dredetianu M. Webware-based asynchronous speaking tasks: a three-item model combining oral interview, web submission and presentation. *Lang Test Asia*. 2013;3(2):98-112.
26. Reed DJ, Cohen AD. Revisiting raters and ratings in oral language assessment. In: Elder C, Brown A, Grove E, Hill K, Iwashita N, Lumley L, *et al.*, editors. *Studies in language testing*. Vol. 11. Experimenting with uncertainty: essays in honour of Alan Davies. Cambridge: Cambridge University Press; 2001. p. 82-96.
27. Stephenson M. Group-based assessments and L2 interactional competence in international student cohorts. *Lang Test*. 2025;42(1):23-45.
28. Taylor L. *Examining speaking: research and practice in assessing second language speaking*. Cambridge: Cambridge University Press; 2011.
29. Thomas RM. *Blending qualitative and quantitative research methods in theses and dissertations*. Thousand Oaks: Corwin Press; 2003.
30. Van Moere A. Validity evidence in a university group oral test. *Lang Test*. 2006;23(4):411-40.
31. Weir CJ. *Language testing and validation: an evidence-based approach*. Basingstoke: Palgrave Macmillan; 2005.
32. Weng J. A systematic review of mobile-assisted peer feedback in EFL speaking: implications for assessment design. *Lang Test Asia*. 2025;15(4):1-17.

How to Cite This Article

Lam N. Ensuring fairness and validity in EFL oral test administration: a socio-cognitive perspective from Vietnamese universities. *Int J Multidiscip Res Growth Eval*. 2025;6(6):54–9. doi:10.54660/IJMRGE.2025.6.6.54-59.

Creative Commons (CC) License

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.