

Research Article

Speaking Assessment Literacy of Iranian EFL Teachers: A Qualitative Study of Their Challenges, Needs, and Strategy Development

Mehrdad Amiri ^{1*}, Zeinab Alizadeh ², Saeedeh Abdollahi Saryazdi ³

1. Assistant Professor, Department of English Language Teaching, Farhangian University, Tehran, Iran.

(Corresponding author) amirimehrdad@cfu.ac.ir

2. MA in TEFL, Ministry of Education; Tehran, Iran, zeinabalizadeh2021@gmail.com

3. M.A. in TEFL, Lecturer at TEFL Research Center, Tehran, Iran, saeedehabdollahisaryazdi@yahoo.com

ARTICLE INFO

Submission History

Received: 2025-07-25

Accepted: 2025-09-15

Keywords

Speaking Assessment Literacy
Strategy Development
Assessment Challenges
Assessment Needs
EFL Teachers

ABSTRACT

This study explores the development of speaking assessment literacy (SAL) among English as a Foreign Language (EFL) teachers in Iran, with particular focus on their strategic practices, encountered challenges, and expressed needs in assessing learners' speaking abilities. Anchored in a grounded theory methodology, the research was initiated through a comprehensive literature review and preliminary interviews with language assessment experts to inform instrument design. Subsequently, semi-structured, in-depth interviews were conducted with some high school EFL teachers. The data were analyzed using open, axial, and selective coding techniques to generate key themes and conceptual categories. Findings revealed that participants recognized speaking assessment as a core instructional responsibility and adopted a variety of strategies, including the use of multiple drafts, provision of assessment rubrics, and both explicit and implicit oral corrective feedback. Additionally, EFL teachers reported designing context-specific rubrics to align with classroom objectives. However, they highlighted some notable challenges they encounter while developing SAL, including students' limited familiarity with rubrics, linguistic underperformance, task-related anxiety, difficulty maintaining assessment objectivity, and insufficient training in self- and peer-assessment practices. Teachers articulated a need for enhanced knowledge of speaking assessment principles and learner-centered assessment approaches. The study underscores the necessity of targeted professional development initiatives to foster EFL teachers' SAL and proposes pedagogical and policy-level interventions to support sustainable assessment practices.

Introduction

Teacher assessment literacy (TAL) has become a cornerstone of effective teaching and learning, particularly in language education. It refers to the ability of educators to design, select, administer, interpret, and use assessment tools to support instructional decisions and promote student learning outcomes (Stiggins, 1991; Weng & Shen, 2022). TAL encompasses knowledge of assessment principles, familiarity with various forms of testing, and the capacity to align assessment practices with curriculum goals and learner needs (Kremmel & Harding, 2020). It enables teachers to collect valid and reliable evidence of learners' performance, diagnose strengths and weaknesses, and adjust instruction accordingly (Jalilzadeh et al., 2023a; Wang et al., 2023). Despite its importance, however, research suggests that many teachers, particularly those in English as a Foreign Language (EFL) contexts, lack sufficient preparation and ongoing support to develop robust assessment literacy (Giraldo, 2018; Vogt & Tsagari, 2014). Within this broader framework, one critical yet underexplored dimension is speaking assessment literacy (SAL), which demands unique theoretical and practical competencies.

As such, assessing oral proficiency requires specialized tools, such as recordings, rubrics, and task-based assessments, as well as interpretive judgment that balances validity, reliability, and authenticity (Ginther, 2012; Fulcher, 2018). SAL, therefore, involves not only an understanding of the theoretical foundations of oral language assessment but also practical expertise in implementing fair and effective assessments in everyday classroom settings (De Jong, 2023; Koizumi, 2022). Yet, in comparison to other domains of TAL, SAL remains significantly

underdeveloped in both research and practice.

This research gap is particularly pronounced in the Iranian EFL context, where speaking instruction is often deprioritized in favor of grammar and reading, and where systemic constraints, such as inadequate teacher training, traditional assessment cultures, and rigid curricula, further inhibit the development of SAL (Farhady, 2022; Pishghadam et al., 2021). While some emerging studies have begun to explore assessment literacy among Iranian EFL teachers, there remains limited empirical insight into how these educators specifically approach the assessment of speaking, what strategies they employ, and what challenges they encounter in doing so (Duque-Aguilar, 2021; Fitri et al., 2021; Idayani & Rugaiyah, 2017). Addressing this gap, the present study investigates how Iranian EFL teachers conceptualize and practice speaking assessment, the contextual and institutional factors shaping these practices, and the forms of support they need to improve their SAL. By shedding light on these dimensions, the study aims to contribute to more targeted teacher education programs, inform assessment policy development, and enhance the overall quality of speaking assessment in EFL classrooms.

Literature Review

Teacher Assessment Literacy (TAL)

The concept of TAL has evolved considerably over the past few decades, emerging as a critical area of investigation in the fields of language education and applied linguistics. TAL refers to the knowledge, competencies, and beliefs that educators hold concerning the design, implementation, interpretation, and ethical use of assessment. As highlighted by Inbar-Lourie (2022), a

teacher's assessment literacy is not merely a technical skill but also involves an understanding of broader socio-political, cultural, and pedagogical dimensions of assessment. According to Fulcher (2018), assessment-literate teachers are those who can design assessment tools aligned with learning objectives, interpret student performance with fairness and validity, and reflect critically on their own assessment practices. Popham (2009) further elaborated on the significance of assessment literacy, contending that without it, teachers may resort to ineffective, unreliable, or even harmful assessment practices.

Speaking Assessment Literacy (SAL)

Within the broader framework of TAL, speaking assessment literacy (SAL) has gained growing prominence. SAL refers to teachers' specialized knowledge and competencies related to the assessment of learners' spoken language. This encompasses familiarity with constructs such as fluency, accuracy, pronunciation, discourse management, pragmatic appropriateness, and interactive competence (Vogt et al., 2024; Taylor, 2013). Moreover, SAL includes the ability to develop or adapt rubrics, apply various scoring methods, provide effective oral and written feedback, and integrate assessment seamlessly into instruction (Moon et al., 2024). However, research indicates that SAL remains a challenging area for many teachers, particularly in EFL contexts, due to limited training opportunities, lack of practical exposure, and insufficient institutional support (Idayani & Rugaiyah, 2017; Fitri et al., 2021).

The importance of developing robust speaking assessment practices is emphasized by Crusan et al. (2016), who argue that inadequately designed assessments can result

in significant negative consequences, including misjudgment of learner abilities, decreased motivation, and erosion of learner confidence. Therefore, it is critical that foreign language (FL) teachers demonstrate a sufficient level of SAL to ensure the accuracy, fairness, and usefulness of their speaking assessments. However, a growing body of research points to persistent gaps in teachers' assessment literacy (Brown & Bailey, 2008; Malone, 2013; Popham, 2009; Scarino, 2013; Sobouti et al., 2023; Stiggins, 1991, 2014; Taylor, 2013). Indeed, both Mertler (2009) and Stiggins (1991) highlight that many teachers report feeling underprepared to assess students' speaking performance effectively and lack confidence in their evaluation practices.

Although assessment literacy has been examined extensively in relation to general classroom assessment and learner achievement (e.g., Jalilzadeh et al., 2023a, 2023b; Jalilzadeh & Coombe, 2023; Tsagari, 2017), relatively few studies have specifically addressed SAL as a distinct domain within teacher assessment literacy. Existing research on SAL remains limited, with notable contributions from scholars such as De Jong (2023), Koizumi (2022), Qian (2008), and Tajeddin et al. (2017, 2018, 2022). This limited attention is particularly concerning given that speaking ability not only serves as a key indicator of communicative competence but also reflects broader processes of second language acquisition (Douglas, 1994).

Empirical research has begun to shed light on the scope and limitations of teachers' SAL. For example, Tajeddin et al. (2018) investigated novice and experienced English teachers' understanding of speaking assessment. Their findings revealed that while both groups possessed knowledge of core evaluation components—such as accuracy,

fluency, pronunciation, interaction, and communicative functions—they differed in their use of specific assessment techniques, including role-plays, monologues, and task-based activities. The study concluded that although both novice and experienced teachers demonstrated awareness of key assessment principles, their practices varied significantly, suggesting the need for more targeted professional development.

Similarly, Qian (2008), in a study conducted in the context of Hong Kong's school-based assessments, reported that many English language teachers lacked the necessary expertise to accurately grade speaking tasks. This deficiency often led to inconsistent scoring and questionable assessment outcomes. Further evidence is provided by Tajeddin et al. (2017), who examined novice teachers' perceptions of SAL in four domains: (a) rating scales for assessing speaking performance, (b) the use of formative versus summative assessment, (c) the roles of teacher, peer, and self-assessment, and (d) the design and implementation of speaking assessment tasks. Their findings indicated a strong preference for analytic rating scales over holistic ones, due to perceived accuracy. However, novice teachers were found to rely predominantly on summative assessments while underutilizing formative approaches. Additionally, their assessment practices were largely shaped by coursebook-driven tasks, with limited use of peer- and self-assessment strategies.

The Iranian context presents additional complexities. According to Rezai (2019), Iranian EFL teachers often struggle with aligning speaking assessment with communicative language teaching (CLT) principles, as traditional exam-oriented approaches continue to dominate educational

settings. Furthermore, studies by Alavi and Kaivanpanah (2007) and Farhady et al. (2010) have emphasized systemic barriers such as limited teacher autonomy, lack of professional development opportunities, and an overemphasis on written exams, all of which negatively affect teachers' ability to assess speaking effectively. Despite growing awareness of these issues, there remains a paucity of research specifically examining how Iranian EFL teachers conceptualize and implement SAL strategies in their classrooms.

Collectively, these findings highlight that while some foundational SAL competencies are present among novice FL teachers, their assessment practices remain constrained and underdeveloped. This underscores the need for systematic, context-sensitive training programs aimed at enhancing teachers' capacity to assess speaking in valid, reliable, and pedagogically meaningful ways.

This study aims to address this gap by adopting a grounded theory approach to uncover the strategies, challenges, and needs related to speaking assessment among EFL teachers in Iran. By doing so, it contributes to the growing body of knowledge on SAL and provides practical implications for teacher education programs, curriculum developers, and educational policymakers.

EFL Teachers' Strategies for Speaking Assessment

Developing effective SAL among EFL teachers requires a thorough understanding of both instructional strategies and their implications for assessment. SAL goes beyond the ability to rate student performances; it encompasses the capacity to design, interpret, and act upon speaking assessments that align with pedagogical objectives. In many EFL classrooms, a variety of speaking instruction

strategies—such as cooperative learning, role-playing, creative tasks, and drilling—are widely used (Melendez, 2014). While these are primarily teaching strategies, they inherently reflect assessment beliefs and shape how teachers evaluate speaking ability. For instance, the use of cooperative tasks suggests an emphasis on interactional competence, a key domain in communicative speaking assessment (Nation & Newton, 2009).

Role-playing activities, which involve students acting out real-life scenarios, provide both instructional and informal assessment opportunities. These activities allow teachers to observe learners' pragmatic competence, fluency, and use of context-appropriate language—components that should be integrated into speaking assessment rubrics (Harmer, 2001; Solcova, 2011; Thornbury, 2005). Similarly, creative tasks reflect an understanding of task-based assessment principles. When learners are engaged in problem-solving or production-oriented activities, they demonstrate authentic language use, which offers rich assessment data for evaluating communicative effectiveness (Solcova, 2011). Drilling, although more form-focused, can also inform diagnostic assessment of pronunciation and accuracy, guiding subsequent instruction and feedback (Thornbury, 2005).

Recent empirical studies further highlight how instructional practices are intertwined with assessment literacy. For example, Amardiaz (2024) identifies description tasks as a useful tool for eliciting spontaneous speech, which contributes to assessing fluency and coherence. Likewise, Husnaini et al. (2024), through a narrative qualitative study in Indonesian high schools, observed that teachers applied a range of speaking activities—including discussions, group work, and

storytelling—not only to teach but also to gauge learners' communicative development. However, the study also revealed significant barriers to effective assessment, such as learners' low proficiency, lack of motivation, fatigue, and insufficient resources. These findings underscore the contextual challenges that shape teachers' assessment practices and point to the need for situated SAL—one that considers the local constraints and affordances in assessment design and implementation.

Crucially, the effectiveness of speaking tasks—both for instruction and assessment—depends on their design. Well-constructed tasks maximize opportunities for authentic language production, a central principle in performance-based assessment (Brown, 2011; Thornbury, 2005). Moreover, tasks should promote meaningful interaction, allowing learners to express opinions, ask questions, and negotiate meaning—skills that align with global frameworks for speaking assessment (e.g., CEFR, TOEFL iBT). Teachers' ability to design such tasks indicates a level of assessment literacy that incorporates understanding of construct validity, task authenticity, and learner engagement. As Brown (2011) and Richards and Rodgers (2002) emphasize, connecting new tasks to learners' existing knowledge facilitates retention, suggesting that assessment should be embedded within familiar, contextually relevant discourse situations.

Furthermore, the dual functions of spoken language—as transactional (focused on information exchange) and interactional (aimed at relationship-building)—highlight the multidimensional nature of speaking assessment (Bailey, 2005; Brown & Yule, 1999; Richards, 2008; Thornbury, 2005). Teachers with high assessment literacy recognize these functions and can distinguish

between them in both instructional and evaluative contexts, designing tasks that align with distinct assessment purposes (e.g., assessing fluency vs. assessing pragmatic appropriateness).

Learning a foreign language is inherently complex, involving the dynamic interplay between teachers, learners, and content. This educational triangle (Díaz Larenas, 2011) shapes how speaking instruction and assessment are conceptualized. Assessment literacy, in this framework, entails the teacher's ability to align content (tasks), pedagogy (strategies), and evaluation (assessment tools) to support student learning. According to Zhang et al. (2022), strategies are structured actions that help learners acquire and retain knowledge—yet they are also mechanisms through which teachers gather evidence of learning. In other words, when teachers facilitate speaking strategies, such as negotiating meaning or using gestures, they are not only supporting communication but also engaging in formative assessment practices that inform their judgments about student progress (De Jong, 2023; Díaz Larenas, 2011).

In light of these considerations, this study aims to investigate the development of speaking assessment literacy among Iranian EFL teachers. It focuses specifically on how teachers conceptualize, apply, and evaluate speaking strategies as part of their assessment practices. Accordingly, the study addresses the following research questions:

1. What strategies do Iranian EFL teachers use to assess L2 speaking?
2. What are the main challenges Iranian EFL teachers face in assessing L2 speaking?
3. What are the perceived needs of Iranian EFL teachers in assessing L2 speaking?

Method

Design of the study

This qualitative study employed a **deductive-inductive research design** to explore Iranian EFL teachers' speaking assessment literacy (SAL) in Iran. In the deductive phase, semi-structured interview questions were developed based on relevant literature in second language assessment and pedagogy, and refined with input from ten TEFL experts affiliated with Farhangian Teacher Education University. In the inductive phase, data were collected from fifty stratified randomly selected in-service EFL teachers through interviews and analyzed using Braun and Clarke's (2021) thematic analysis to identify emerging themes. The study ensured contextual and theoretical validity, combining expert insights with classroom-based perspectives, and maintained methodological rigor through, member checking, and ethical research practices.

Participants and Setting

This qualitative study was involved two distinct groups of participants, namely expert teacher educators and in-service EFL teachers in Iran. The first group held doctoral degrees in Teaching English as a Foreign Language (TEFL) and were affiliated with Farhangian Teacher Education University in Tehran, Iran. We purposefully selected these individuals based on their scholarly expertise in second language acquisition (SLA), language pedagogy, and assessment. Their role in the study was to contribute to the initial development and refinement of the semi-structured interview guide, ensuring that the protocol was theoretically grounded and contextually relevant. Through one-on-one interviews, the expert participants provided insight into strategic approaches to L2

speaking development and assessment, which subsequently informed the content and structure of the interview questions used with the main participant group.

The second group comprised fifty in-service EFL teachers currently employed in public secondary schools across various educational districts in Tehran. A stratified random sampling technique was employed to ensure the inclusion of teachers from diverse institutional and geographical contexts within the city. Demographically, the sample included 21 male and 29 female teachers, representing a range of academic qualifications: six held doctoral degrees in TEFL, eleven held Master's degrees in either TEFL or English Literature, and thirty-three held Bachelor's degrees in TEFL. Participants were selected from an initial pool of 150 teachers who met the inclusion criteria (minimum three years of teaching experience and current employment in an EFL context), from which the final sample was drawn using a computer-generated randomization process.

The in-service teachers participated in semi-structured interviews designed to elicit in-depth accounts of the strategies they employ in assessing learners' L2 speaking abilities, the practical and institutional challenges they encounter in doing so, and the professional development needs they perceive in relation to speaking assessment literacy (SAL). These interviews aimed to explore the teachers' operational knowledge, experiential insights, and attitudinal orientations toward speaking assessment in their classrooms. All participants received an information sheet detailing the purpose and procedures of the study and provided written informed consent. Confidentiality was maintained throughout the research process; pseudonyms were used in all

transcriptions and reporting to protect participant identity.

The selection of both expert and practitioner participants was intended to ensure a comprehensive understanding of speaking assessment literacy, combining theoretical expertise with practical classroom perspectives. This two-tiered participant design facilitated the development of an empirically grounded and contextually valid research instrument while capturing a broad range of authentic assessment practices and beliefs among Iranian EFL teachers.

Instruments and data collection

Data for this study were collected in two sequential phases over a period of three months, from January to March 2025, using semi-structured interviews as the primary qualitative data collection method. This approach was selected to allow for in-depth exploration of participants' perspectives, professional experiences, and assessment practices related to L2 speaking. The first phase involved conducting interviews with ten expert teacher educators, whose input was instrumental in shaping the content validity of the main interview guide. Drawing on their disciplinary knowledge in second language acquisition, language assessment, and teacher education, these participants were asked to reflect on essential strategies for L2 speaking instruction and assessment, as well as current gaps in EFL teachers' assessment literacy. The interviews were conducted individually, either face-to-face at Farhangian Teacher Education University or virtually via secure video conferencing platforms, depending on participant preference and availability. Each interview lasted between 45 and 60 minutes. The sessions were audio-recorded with participant consent and transcribed verbatim

for subsequent thematic coding. Insights derived from this phase informed the construction of the final interview guide used in the second phase. In the second phase, semi-structured interviews were conducted with 50 in-service EFL teachers working in public high schools across Tehran. The interview protocol consisted of open-ended questions organized into three thematic domains aligned with the study's research questions: (1) strategies used in assessing L2 speaking, (2) challenges experienced during speaking assessment, and (3) perceived professional needs related to speaking assessment literacy. Probing questions were used to elicit specific examples, rationales, and reflections on pedagogical and institutional influences shaping their assessment practices (see Appendix 1 for more information).

Each interview was conducted in a quiet location mutually agreed upon by the researcher and the participant, typically within the school premises or via an encrypted video call for those unable to attend in person. Interviews ranged from 40 to 70 minutes in length. All interviews were conducted in Persian to ensure linguistic comfort and depth of expression, and were subsequently translated into English for analysis, with back-translation employed to preserve meaning accuracy.

All interviews were digitally recorded using secure, password-protected devices. Participants were assured of the confidentiality of their responses, and pseudonyms were assigned during transcription to protect individual identities. Informed consent was obtained from each participant prior to the interview, with the assurance that their participation was voluntary and that they could withdraw at any point without any consequences. The study was approved by the

Research Ethics Committee of [insert university name], and all procedures adhered to institutional and international ethical guidelines for research involving human subjects.

This two-phased data collection strategy ensured that the final interview guide was theoretically grounded and contextually tailored, while also enabling rich, experience-based data collection from in-service teachers actively engaged in classroom-based speaking assessment. The triangulation of expert and practitioner insights further enhanced the credibility and comprehensiveness of the data set.

Data Analysis

The analysis of the qualitative data obtained through semi-structured interviews followed a rigorous, iterative process grounded in thematic analysis (Braun & Clarke, 2021). This method was selected for its flexibility in identifying, analyzing, and reporting patterns (themes) within rich qualitative data while allowing the researcher to interpret various aspects of participants' SAL.

All interviews were transcribed verbatim in the original language (Persian) immediately following each interview session to preserve the authenticity and nuance of participants' responses. The transcriptions were reviewed alongside audio recordings for accuracy. English translations were conducted by the lead researcher and checked by a bilingual language education expert using back-translation procedures to ensure conceptual equivalence. Transcripts were then imported into MAXQDA 2022, a qualitative data analysis software, to facilitate systematic coding and organization.

Initial coding was conducted using an inductive approach, allowing codes to emerge

organically from the data rather than being predetermined by an external framework. Each transcript was read multiple times to ensure familiarity with the content, and preliminary codes were generated line-by-line to capture meaningful units of data relevant to the study's focus: speaking assessment strategies, challenges, and professional development needs.

Following initial coding, the researcher organized the codes into broader thematic categories using constant comparative analysis (Glaser & Strauss, 1967). Codes were clustered based on similarities, differences, and relationships, and then synthesized into candidate themes. These themes were iteratively refined, collapsed, or expanded as needed to ensure distinctiveness and analytical depth. This process yielded three overarching thematic categories corresponding to the study's research questions: (a) strategic approaches to L2 speaking assessment (e.g., performance-based tasks, observational assessment, peer-assessment, and improvisational speaking tasks), (b) contextual and pedagogical challenges (e.g., limited time, high student numbers, lack of standard rubrics, and test-related anxiety), and (c) teachers' professional needs and aspirations (e.g., training in rating consistency, exposure to communicative assessment models, and clarity in scoring criteria).

Each theme was supported by direct quotations from participants to preserve voice authenticity and provide evidence for interpretation. Disconfirming evidence and outlier responses were also analyzed to enhance the credibility and nuance of thematic interpretation.

To ensure the trustworthiness of the data analysis, member checking was employed. For this purpose, a subset of participants (n=8) was

provided with synthesized summaries of their interview responses and the emerging themes to confirm the accuracy of interpretations. Moreover, dependability was established through the use of an audit trail, including documentation of coding decisions, reflective memos, and data triangulation between expert and practitioner perspectives.

Overall, the thematic analysis allowed for a rich, contextually grounded understanding of Iranian EFL teachers' speaking assessment literacy, shaped by their instructional realities, theoretical awareness, and institutional constraints.

Results

This section presents the key findings from interviews conducted with 50 Iranian EFL teachers. Thematic analysis was applied to identify recurring patterns, strategies, challenges, and needs related to speaking assessment literacy (SAL). The findings are organized into three major themes: (1) strategies used by EFL teachers in assessing speaking, (2) challenges faced in implementing effective speaking assessment, and (3) perceived needs for developing assessment literacy. Direct quotes from participants are included to provide rich, illustrative evidence for each theme.

Strategies used to develop speaking assessment

Analysis revealed a range of strategies employed by Iranian EFL teachers to assess students' speaking performance. These strategies reflected varying levels of assessment literacy and engagement with best practices. Many participants expressed a sense of professional responsibility toward assessing speaking skills. One teacher remarked: *"Assessing speaking is not an option—it's part*

of my job as a language teacher, even though it's challenging." (Participant 12). While some saw speaking assessment as engaging and essential, others expressed frustration: "Sometimes I feel stuck; I don't always know how to assess fairly, especially when students have very different levels." (Participant 27). This divergence highlights the spectrum of attitudes that influence strategy implementation.

Teachers also reported various degrees of rubric use, with some creating their own assessment tools: "I try to make my own rubrics so that students know what I expect. But I don't always have time to make one for every task." (Participant 5). Others admitted to not using rubrics consistently: "I rarely share rubrics. I just listen and give general comments, but I know that's not enough." (Participant 33). Feedback was widely recognized as a key part of assessment, though the mode and frequency varied: "I usually

discuss their errors right after they finish. It helps them connect the mistake with the context." (Participant 18). Informal strategies such as oral correction, delayed feedback, or peer observation were commonly employed. Teachers sometimes avoided correction to preserve fluency or reduce anxiety. "I ignore minor errors unless they cause confusion. Confidence is more important at times." (Participant 22)

There was inconsistency in how assessment criteria were shared and explained to learners. Some teachers emphasized making expectations clear: "Before every speaking task, I explain what I'm looking for—fluency, vocabulary use, grammar, coherence." (Participant 41). Others admitted that while rubrics were used, learners often did not understand or apply them effectively: "I think students don't really know how to use the rubric. I tell them to look at it, but I don't train them on it." (Participant 17).

Table 1

Strategies Used by Iranian EFL Teachers in Speaking Assessment

Subtheme	Key Insights
Attitudes Toward Speaking Assessment	Mixed attitudes: Some viewed it as essential; others expressed uncertainty or stress.
Use of Rubrics and Feedback	Inconsistent rubric use; feedback varied in type and timing; informal methods were common.
Student Engagement with Criteria	Criteria were not always explained clearly; students often misunderstood or underutilized rubrics.

Challenges faced in speaking assessment

Thematic analysis of the interview data uncovered six major themes capturing the core challenges teachers face in assessing L2 speaking. Teachers struggled to balance multiple assessment factors such as fluency, grammar, coherence, and pronunciation: "Sometimes students speak fluently but make grammar mistakes. Should I focus on fluency or correctness?" (Participant 6). This lack of

clarity led to inconsistency in judgments and uncertainty about fairness.

Some teachers reported relying on a narrow range of assessment techniques, often determined by time or institutional constraints: "Most of the time we just do role-plays or interviews. I wish we had more variety, but the syllabus is tight." (Participant 29). Few participants were familiar with dynamic or formative assessment methods.

Participants expressed varying approaches to dealing with learner errors. While some favored oral corrective feedback (OCF), others were more passive: *“I let students finish, then I note down the big mistakes and talk to them later. I don’t interrupt unless it’s serious.”* (Participant 14). Peer correction and self-assessment were rarely integrated in a structured manner.

A major theme was the limited application of assessment literacy knowledge. Some teachers were aware of practices such as peer and self-assessment but lacked confidence in applying them: *“I know about peer-assessment, but honestly, I’ve never tried it. I’m not sure how to guide it properly.”* (Participant 35).

Many teachers emphasized the influence of affective and sociocultural factors: *“Some students freeze when asked to speak. Their anxiety makes it hard to assess their true level.”* (Participant 10). Creating a supportive classroom environment was viewed as essential: *“I try to lower their stress by being friendly and not pointing out every mistake.”* (Participant

A final challenge emerged around teachers’ own proficiency and assessment literacy: *“How can I judge their speaking when my own speaking isn’t perfect?”* (Participant 3). Others reflected on their lack of formal training in assessment: *“We never had courses on how to assess speaking. We just learn by doing—and sometimes guessing.”* (Participant 46).

Table 2
Challenges Faced in Assessing Speaking

Subtheme	Key Insights
Complex Criteria for Evaluation	Difficulty balancing accuracy and fluency; lack of clarity on evaluation priorities.
Limited Assessment Techniques	Over-reliance on interviews/role-plays; limited exposure to varied assessment techniques.
Handling Learner Errors	Mixed approaches to error correction; structured peer/self-assessment was rare.
Gaps in Implementing Assessment Knowledge	Teachers lacked confidence or experience in applying known assessment practices.
Psychological and Sociocultural Factors	Anxiety and classroom dynamics affected assessment accuracy; emotional support seen as important.
Teachers’ Speaking Ability and SAL	Self-doubt in language proficiency and lack of formal training impeded assessment practices.

Perceived needs to improve SAL

Participants identified both theoretical and practical needs to enhance their competence in assessing speaking. There was a widespread desire for more structured training: *“I’d love to attend workshops on speaking assessment—how to use rubrics, how to evaluate fluency fairly.”* (Participant 15). Reading research, watching videos, and participating in professional communities were mentioned as helpful: *“I’ve started reading articles on*

assessment, but I need more guidance to apply them.” (Participant 39).

Teachers called for clearer, learner-friendly rubrics and more training in how to introduce them: *“It would be great to have sample rubrics tailored to each level—students would know what to expect.”* (Participant 26). They also emphasized the need to consider learners’ preferences and individual differences: *“Each student is different. We need to adjust our assessment and feedback accordingly.”*

(Participant 42). Overall, participants advocated for more continuous professional development opportunities and greater institutional support.

Table 3

Perceived Needs to Improve Speaking Assessment Literacy

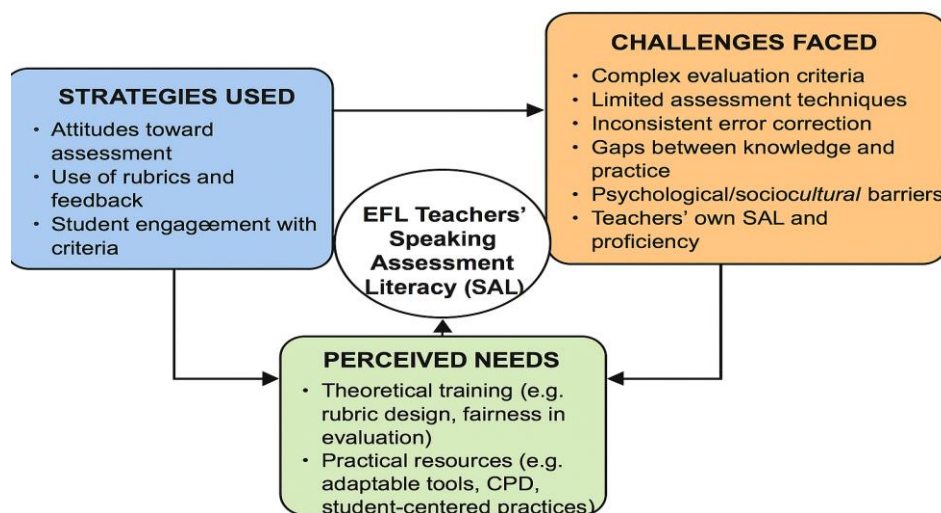
Subtheme	Key Insights
Theoretical Knowledge	Teachers wanted structured training, academic resources, and peer learning opportunities.
Practical Tools and Strategies	Desire for level-appropriate rubrics, adaptable strategies, and responsive assessment tools.

The findings underscore the complexities Iranian EFL teachers face in assessing speaking. While many teachers are committed and resourceful, they often lack formal training, clear guidelines, and support systems to implement effective and fair speaking assessments. Addressing their professional

development needs—particularly in rubric design, feedback strategies, and affective awareness—is essential for improving their speaking assessment literacy. Figure 1 demonstrates EFL teachers' perceptions of overall speaking assessment literacy (SAL).

Figure 1

EFL Teachers' Speaking Assessment Literacy (SAL)



Discussion and Conclusion

This study aimed to explore Iranian EFL high school teachers' strategies for assessing speaking skills and to examine the extent to which these strategies reflect their Speaking Assessment Literacy (SAL). Overall, the findings suggest that while many teachers possess a sound understanding of theoretical principles related to speaking assessment,

their ability to implement these strategies effectively in the classroom is constrained by contextual and systemic factors.

The data revealed that teachers generally value speaking assessment and view it as an integral aspect of their professional role. Many participants indicated that they attempt to support their learners by using formative strategies such as multiple drafts, provision of

rubrics or assessment criteria, and various forms of oral corrective feedback. These practices reflect a commitment to student-centered assessment and are in line with established principles of effective language assessment, suggesting a degree of strategic awareness among teachers.

In some cases, teachers reported developing their own rubrics or modifying existing ones to better suit the classroom context, indicating a degree of autonomy and engagement with assessment design. They also commonly mentioned the use of both explicit and indirect oral corrective feedback to scaffold learners' development, supporting the idea that assessment is not separate from instruction but intertwined with ongoing language learning. These practices resonate with the findings of Tseng et al. (2006), who emphasized the reciprocal relationship between teaching and assessment strategies in shaping learner outcomes.

However, despite this conceptual understanding, teachers often reported that they struggle to implement these strategies effectively. One of the most prominent challenges mentioned was the lack of time to conduct thorough speaking assessments due to overloaded curricula and large class sizes. Teachers also pointed to systemic constraints, including a school system that gives limited attention to English language education, reducing the perceived value and support for developing learners' speaking skills. Some participants also noted that sociopolitical factors, such as national priorities in curriculum design, hinder their efforts to focus on communicative language teaching and assessment.

This gap between knowledge and practice raises concerns about the adequacy of teacher preparation programs in addressing the

practical demands of speaking assessment. Although most participants had completed courses in language testing and evaluation during their undergraduate or graduate studies, these courses often emphasized standardized testing and lacked sufficient focus on alternative or classroom-based assessment strategies. As Mertler (2009) pointed out, the shift from measurement to assessment in education requires updated training that reflects current pedagogical trends.

Recent research supports the need to move beyond traditional assessment models. For example, competency-based (Gallardo, 2020) and performance-based assessment (Tosuncuoğlu, 2018) approaches require teachers to be able to design, implement, and interpret student-centered tasks that reflect real-world communication. Unfortunately, these concepts are not yet fully integrated into teacher education in many contexts, including Iran. As a result, many teachers feel unprepared to apply such models in their classrooms.

Furthermore, the findings support Amiri and Birjandi's (2015) argument that SAL should include interlanguage pragmatic knowledge. Since speaking often involves the use of language in contextually appropriate ways, teachers must also be equipped to evaluate pragmatic competence as part of their assessment practice. Without this dimension, speaking assessments risk neglecting the sociolinguistic aspects of communication.

In light of these findings, there is a clear need for sustained professional development programs that go beyond general knowledge of assessment and address the specific skills required for speaking assessment in communicative and contextualized settings. By supporting teachers in implementing

assessment strategies aligned with their pedagogical values, such programs could contribute to more meaningful and effective speaking instruction in Iranian high schools.

Limitation and Suggestion for Further Research

A key limitation of the present study lies in its reliance on self-reported data obtained through semi-structured interviews. While interviews allowed for rich, in-depth exploration of teachers' beliefs, experiences, and perceived practices regarding speaking assessment, they are inherently subject to limitations such as social desirability bias and selective memory. Participants may have presented idealized accounts of their assessment behaviors or may have unintentionally omitted aspects of their actual classroom practice. As a result, the findings may not fully reflect the complexities and contextual nuances of real-time assessment decisions.

To address this limitation and enrich our understanding of speaking assessment literacy (SAL) among EFL teachers, future research is encouraged to adopt a more comprehensive, mixed-methods approach. This could include classroom observations, stimulated recall interviews, or video-based reflection tasks to triangulate self-reported data with observable assessment practices. Moreover, longitudinal studies that follow teachers over time—particularly in the context of professional development programs focused on SAL—could provide valuable insights into how assessment knowledge and practices evolve. Investigating the effectiveness of interventions such as rubric training, feedback workshops, or peer-assessment modules would contribute significantly to the design of evidence-based teacher education initiatives.

References

- Amardiaz, R. (2024). Students' interest in speaking instruction through describing object method at junior high school. *English Franca Academic Journal of English Language and Education*, 87(1), 243-254. <http://202.162.210.182/index.php/english>
- Amiri, M. (2020). *Second language research methods*. Arsa Publications. <https://www.gisoom.com/book/11572505/%D%A%A9%D8%AA%D8%A7%D8%A8-Second-language-research-methods/>
- Amiri, M., & Birjandi, P. (2015). Reliability and content validity of a comprehensive discourse completion test of inter-language pragmatics for academic situations. *Sino-US English Teaching*, 12(9), 654-659. <https://doi.org/10.17265/1539-8072/2015.09.002>
- Amiri, M., Nezakatgoo, B., & Jalilzadeh, K. (2024). Developing and validating a pedagogical content knowledge (PCK) scale for the TEFL teacher education program in the Iranian context. *Journal of Research in Curriculum Studies*, 1(1), 1-15. <https://doi.org/10.48310/jcdr.1403.3722>
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.
- Bailey, K. M. (2005). *Practical English language teaching*. McGraw Hill.
- Best, J. W., & Kahn, J. V. (2016). *Research in education*. Pearson Education India.
- Brown, G. T. L. (2011). Teachers' conceptions of assessment: Comparing primary and secondary teachers in New Zealand. *Assessment Matters*, 3, 45-70.
- Brown, G., & Yule, G. (1999). *Teaching the spoken language*. Cambridge University Press.
- Brown, J. D., & Bailey, K. M. (2008). Language testing courses: What are they in 2007? *Language Testing*, 25(3), 349-383.
- Burns, A. (2012). A holistic approach to teaching speaking in the language classroom. In M. Olofsson (Ed.), *Symposium 2012* (pp. 165-178). Liber.
- Bygate, M. (2001). Speaking. In R. Carter & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp. 14-20). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667206.003>.
- Coombe, C., & Hubley, N. (2011). *Fundamentals of language assessment*. Mexico Documents.
- Correia, R. C. (2024). The art and science of assessment—Connecting classroom-based

- speaking assessments to teachers' practice. *M. Carmo, Education Applications & Developments IX*, 49-66.
- Creswell, J. W. (2021). *A concise introduction to mixed methods research*. SAGE publications.
- Creswell, J. W., & Plano Clark, V. L. (2023). Revisiting mixed methods research designs twenty years later. In N. P. Cheryl (Ed.), *Handbook of mixed methods research designs* (pp. 21-36). SAGE Publications Ltd.
- 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>.
- Cutcliffe, J. R., & McKenna, H. P. (1999). Establishing the credibility of qualitative research findings: the plot thickens. *Journal of advanced nursing*, 30(2), 374-380.
- De Jong, N. H. (2023). Assessing second language speaking proficiency. *Annual Review of Linguistics*, 9(1), 541-560.
- Díaz Larenas, C. (2011). Exploring knowledge of English-speaking strategies in 8th and 12th graders. *Profile Issues in Teachers Professional Development*, 13(2), 85-98.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.
- Douglas, D. (1994). Quantity and quality in speaking test performance. *Language Testing*, 11(2), 125-144.
- Duque-Aguilar, J. F. (2021). Teachers' assessment approaches regarding EFL students' speaking skill. *Profile Issues in Teachers Professional Development*, 23(1), 161-177.
- Farhady, H. (2022). Language testing and assessment in COVID-19 pandemic crisis. In K. Sadeghi (Ed.), *Technology-Assisted language assessment in diverse contexts* (pp. 54-68). Routledge.
- Fitri, A. Y., Hati, G. M., & Lubis, A. A. (2021). English teachers' strategies in assessing eighth graders' speaking skill. *Wacana: Jurnal Penelitian Bahasa, Sastra dan Pengajaran*, 19(2), 1-6.
- Fulcher, G. (2018). Assessing spoken production. In J. I. Lontos, TESOL International Association, & M. Del-liCarpini (Eds.), *The TESOL encyclopedia of English language teaching* (pp. 1-6). <https://doi.org/10.1002/9781118784235.eelt0364>.
- Gallardo, K. (2020). Competency-based assessment and the use of performance-based evaluation rubrics in higher education: Challenges towards the next decade. *Problems of Education in the 21st Century*, 78(1), 61-79.
- Ginther, A. (2012). Assessment of speaking. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Blackwell. <https://doi.org/10.1002/9781405198431.wbeal0052>.
- Green, A. (2013). *Exploring language assessment and testing: Language in action*. Routledge. <https://doi.org/10.4324/9781315889627>.
- Harmer, J. (2001). *The practice of English language teaching* (3rd ed.). Longman.
- Husnaini, H., Syam, A. T., Niken, N., & Soares, A. (2024). Teachers' strategies in teaching speaking at high schools: Obstacles and challenges. *Linguistics Initiative*, 4(2), 243-262.
- Hutagaol, N. N., & Oktora, M. (2022). EYL teacher's strategies in conducting speaking assessments. *Register: Journal of English Language Teaching of FBS-Unimed*, 11(1), 1-17.
- Idayani, A., & Rugaiyah, R. (2017). An analysis of teachers' strategies in conducting speaking assessments at MTSN Andalan Pekanbaru. *J-SHMIC: Journal of English for Academic*, 4(1), 1-14.
- Inbar-Lourie, O. (2022). EMI programs and formative assessment: Implications for the assessment literacy of content lecturers. *Journal of English-Medium Instruction*, 1(2), 204-231. <https://doi.org/10.1075/jemi.21014.inb>
- Jalilzadeh, K., & Coombe, C. (2023). Constraints in employing learning-oriented assessment in EFL classrooms: Teachers' perceptions. *Language Testing in Asia*, 13(1), 7-21.
- Jalilzadeh, K., Alavi, S. M., & Siyyari, M. (2023a). A qualitative study of language assessment literacy of Iranian EFL teachers with TEFL and Non-TEFL background and challenges in their students' assessment. *Language and Translation*, 13(4), 93-109.
- Jalilzadeh, K., Coombe, C., Nezakatgoo, B., & Bagherpour, N. (2023b). Demystifying the challenges of e-assessment in Turkish preparatory language schools. *Computer-Assisted Language Learning Electronic Journal*, 24(2), 129-155.
- Koizumi, R. (2022). L2 speaking assessment in secondary school classrooms in Japan. *Language Assessment Quarterly*, 19(2), 142-161.
- Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups:

- Developing the language assessment literacy survey. *Language Assessment Quarterly*, 17, 100-120.
<https://doi.org/10.1080/15434303.2019.1674855>
- Mackey, A., & Gass, S. M. (2016). *Second language research: Methodology and design*. Routledge.
- Malone, M.E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329-344.
- Mauranen, A. (2006). Spoken discourse, academics, and global English: A corpus perspective. In R. Hughes (Ed.), *Spoken English, TESOL, and applied linguistics: Challenges for theory and practice* (pp. 143-158). Palgrave Macmillan.
https://doi.org/10.1057/9780230584587_7.
- Maxwell, J. A. (2022). Interactive approaches to qualitative research design. In U. Flick (Ed.), *The Sage handbook of qualitative research design* (pp. 41-45). SAGE publications.
- Melendez, R. A. M., Zavala, G. G. Q., & Mendez, R. F. (2014). Teaching speaking strategies to beginners. *Proceedings of 1st international scientific forum, ISF 2013*, 12-14 December 2013, Tirana, Albania (pp. 458-554). Vitrina University Press.
- Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving schools*, 12(2), 101-113.
- Moon, J., Lee, D., Choi, G. W., Seo, J., Do, J., & Lim, T. (2024). Learning analytics in seamless learning environments: a systematic review. *Interactive Learning Environments*, 32(7), 3208-3225.
<https://doi.org/10.1080/10494820.2023.2170422>
- Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL Listening and Speaking*. Routledge.
- Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, 84, 128-138.
- Pishghadam, R., Derakhshan, A., Jajarmi, H., Tabatabaee Farani, S., & Shayesteh, S. (2021). Examining the role of teachers' stroking behaviors in EFL learners' active/passive motivation and teacher success. *Frontiers in Psychology*, 12, 707314.
- Polat, M. (2020). Investigating the backwash effect of higher education exam (YGS) on university students' attitudes. *International Journal of Psychology and Educational Studies*, 7(3), 152-163.
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 48(1), 4-11.
- Qian, D. D. (2008). English language assessment in Hong Kong: A survey of practices, developments, and issues. *Language Testing*, 25(1), 85-110.
- Richards, J. C. (2008). *Teaching listening and speaking, from theory to practice*. Cambridge: Cambridge University Press.
- Richards, J., & Rodgers, T. (2002). Approaches and methods in language teaching. *Korea TESOL Journal*, 5(1), 161-165.
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309-327.
- Sobouti, F., Gharagozloo, N., & Rahimi, A. (2023). A study on the assessment literacy of novice and experienced ESP instructors with TEFL and non-TEFL backgrounds in Iranian academic context. *Critical Literary Studies*, 5(1), 177-215.
- Solcova, P. (2011). *Teaching Speaking Skills* [Unpublished master's thesis], Masaryk University.
- Stiggins, R.J. (1991). Assessment literacy. *The Phi Delta Kappan*, 72(7), 534-539.
- Stiggins, R.J. (2014). *Revolutionize assessment: Empower students, inspire learning*. Corwin Press.
- Tajeddin, Z., Alemi, M., & Yousefi, L. (2017). *Novice EFL teachers' perceptions of speaking assessment literacy* [unpublished masters' thesis], IKIU Annual Conference on Recent Research in English Language Teaching.
- Tajeddin, Z., Alemi, M., & Yasaei, H. (2018). Classroom assessment literacy for speaking: Exploring novice and experienced English language teachers' knowledge and practice. *Iranian Journal of Language Teaching Research*, 6(3), 57-77.
- Tajeddin, Z., Khatib, M., & Mahdavi, M. (2022). Critical language assessment literacy of EFL teachers: Scale construction and validation. *Language Testing*, 39(1), 23-31.
<https://doi.org/10.1177/02655322211057040>
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403-412.
<https://doi.org/10.1177/0265532213480338>

- Thornbury, S. (2005). *How to teach speaking*. Pearson Educational Limited.
- Tosuncuoğlu, I. (2018). Importance of assessment in ELT. *Journal of Education and Training Studies*, 6(9), 163-167.
- Tsagari, D. (2021). Gauging the assessment literacy levels of English language teachers in Norway. *European Journal of Applied Linguistics*, 10, 161-192.
- Tseng, W. T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics*, 27(1), 78-102.
- Vogt, K., Bøhn, H., & Tsagari, D. (2024). Language assessment literacy. *Language Teaching*, 57(3), 325-340. <https://doi.org/10.1017/S0261444824000090>
- Wang, X., Zuo, J., Liu, F., & Sun, Z. (2023). A scientometric review of research trends in language assessment literacy. *Education Sciences*, 13(2), 190-201. <https://doi.org/10.3390/educsci13020190>
- Webb, N. (2002, April). Assessment literacy in a standards-based urban education setting. In F. Davison (Ed.), *Annual meeting of the American educational research association* (pp. 1-20). Arnolds.
- Weng, F., & Shen, B. (2022). Language assessment literacy of teachers. *Frontiers in Psychology*, 13, 864582.
- Xu, N. (2020). The backwash effect of TEM4 on comprehensive English teaching. *International Journal of Social Science and Education Research*, 3(6), 37-43.
- Zhang, W., Zhao, M., & Zhu, Y. (2022). Understanding individual differences in metacognitive strategy use, task demand, and performance in integrated L2 speaking assessment tasks. *Frontiers in Psychology*, 13, 876208.