



EFL Classroom Assessment: Insight and Narratives from Teacher Trainers

Adel Dastgoshadeh

Department of English, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran.

adeldastgoshadeh@mail.com

Hossein Farhadi

ELT Department, Yeditepe University, Istanbul, Turkey.

farhady@gmail.com

Kaveh Jalilzadeh 

Department of Language and Literature, Islamic Azad University, Science and Research Branch, Tehran, Iran.

kaveh.j@iuc.edu.tr

Raheleh Khosravi  (Corresponding Author)

MA in TEFL, University of Kurdistan, Sanandaj, Iran.

rahelekhosravi52@gmail.com

ARTICLE INFO:

Received date:

2025.08.02

Accepted date:

2025.10.15

Print ISSN: 2251-7995

Online ISSN: 2676-6876

Keywords:

Language Assessment Literacy, EFL Teacher Trainers, Teacher Education, Classroom Assessment.



Abstract

Language Assessment Literacy has emerged as a pivotal competence for EFL teacher trainers in guiding the transition from traditional summative testing toward formative, learner-centered assessment practices. This qualitative study investigates how teacher trainers conceptualize and enact assessment in their classrooms, focusing on the role of LAL in enabling pedagogical change. Semi-structured interviews with fourteen EFL teacher trainers were analyzed inductively using MAXQDA, yielding 114 codes categorized into themes related to prior assessment experiences, influential contextual factors, and perceptions of ideal assessment quality. Findings reveal that trainers' current practices are strongly shaped by their own histories as learners, institutional resources, and commitments to meeting students' affective and cognitive needs. While summative methods remain common due to systemic and workload constraints, trainers identify formative and diagnostic assessment as essential for fostering learning, student motivation, and self-concept. Recommendations highlight the need for flexible institutional policies, collaborative professional development workshops, and integration of technology to support formative approaches. This study concludes that strengthening LAL among teacher trainers is critical not only for refining assessment tools and processes but also for driving the broader transformation of EFL assessment culture toward equitable, and pedagogically aligned practices.

Citation: Dastgoshadeh, A.; Farhadi, H.; Jalilzadeh, K. & Khosravi, R. (2025). EFL Classroom Assessment: Insight and Narratives from Teacher Trainers. *Journal of English Language Teaching and Learning*, 17 (36), 99-118. DOI:10.22034/elt.2025.68448.2810

Introduction

In recent decades, there has been a surge of scholarly interest in assessment literacy (AL) within the educational assessment arena, followed by language assessment literacy (LAL) in the field of language education. A broad definition of AL, in the latter field, though not explicitly focusing on language, is provided by Fulcher (2012). It involves the knowledge, skills, and abilities necessary for designing, developing, maintaining, and evaluating large-scale standardized and/or classroom-based tests. Taylor (2013) extends the definition of Language Assessment Literacy (LAL), suggesting that different groups involved in language assessment—such as teachers, test designers, and administrators—are likely to possess different levels of knowledge, skills, and experience. She argues that LAL goes beyond simply knowing how to design or score a test; it includes understanding the theoretical foundation of assessment, having technical expertise, being familiar with main concepts, and applying effective teaching strategies. It also involves being aware of sociocultural influences, local practices, and even personal beliefs or attitudes that can shape how assessment is carried out and interpreted. To illustrate this, Taylor outlined distinct LAL profiles for four groups of stakeholders: professional language testers, test writers, classroom teachers, and university administrators. Building on her work, Baker and Riches (2018) introduced a more nuanced model by incorporating Eraut's (2004) framework of professional expertise. Their version underscores self-awareness, working collaboratively, and recognizing how deeply sociocultural values influence assessment. These advancements helped bring attention to LAL is not a fixed set of skills, but an evolving, reflective, and context-sensitive process.

Another model addressing both the internal and external dimensions of teachers' assessment knowledge is known as Xu and Brown's (2016) Teacher Assessment Literacy in Practice framework (TALiP). Drawing from a range of existing studies on assessment literacy, Xu and Brown (2016) categorized the concept into six key elements: the way teachers view assessment, how they understand and interpret it, the influence of their school environment and broader cultural context, how they actually apply assessment, how they continue to grow and learn in this area, and how teachers form and reshape their identity in assessment.

LAL extends its relevance to encompass all stakeholders, including both those involved in teaching, such as language educators, course leaders, and teacher educators, and non-teaching entities, such as students, parents, program managers, admission officers, student services/advisors, ministerial staff, and policymakers (e.g., Inbar-Lourie, 2017; Taylor, 2009). Despite this multidimensionality of LAL, it is noteworthy that the existing literature predominantly emphasizes on the LAL of educators, as highlighted by Harding and Kremmel (2016). Consequently, there arises a necessity to reframe the scope of LAL to broaden the focus beyond the language and literacy competencies of educators.

However, some researchers point out that not everyone involved in assessment needs the same depth of knowledge or skill. Various groups—like students, teachers, policymakers, and test designers—each play different roles, so their understanding of assessment should reflect that. For instance, Pill and Harding (2013) argue for thinking about assessment literacy as a scale, with five levels: starting from no knowledge (illiteracy), then moving through basic

awareness (nominal), practical understanding (functional), hands-on use (procedural), and finally, deep, critical insight (conceptual). Together, leading to a complex model of what it means to be assessment literate. They suggest that for stakeholders such as policymakers, functional literacy might be adequate, while language testers are expected to possess knowledge of other dimensions of LAL. Shahid Abrar-ul-Hassan and Nassaji (2023) contend that given the ongoing epistemological fluidity in language education, there is a need to reconsider the LAL framework through a "rescoping lens." This lens entails a comprehensive examination of language assessment practices, assessment outcome interpretations, and the utilization of assessment across and beyond educational realms. The rescoping lens signifies a paradigmatic shift in how language assessment is understood and implemented. This shift is underpinned by two fundamental factors: criticality and impact. Through a social constructivist perspective, both criticality and impact are dynamically constructed and reconstructed over time in various social contexts within and beyond educational settings.

Due to the significant role that teachers play in the assessment process, a substantial portion of LAL research has revolved around language teachers, offering invaluable insights into their needs, challenges, and competencies in assessment within language classrooms (Dastgoshadeh & Jalilzadeh, 2024; Hasselgreen et al., 2004; Lam, 2015; Gan & Lam, 2022; Tsagari & Vogt, 2017). Nonetheless, limited studies exist on exploring how language assessment literacy develops among different groups involved in the process like teachers, students, and policymakers. For instance, Gan & Lam (2022) contend that in addition to learners, other stakeholders in LAL studies including policymakers, language testers, admission officers, test developers, teacher educators, and administrators are also underrepresented. More specifically, the literature underscores the dearth of studies explicitly focusing on EFL assessment experts, despite broader discussions on teacher education and language assessment practices. Operating at the confluence of theoretical principles and practical applications, these professionals play a pivotal role in designing, implementing, and innovating language assessment practices. Therefore, recognizing the unique experiences, practices, and literacy of EFL assessment experts becomes imperative for addressing the intricacies and challenges embedded in language assessment. To address the issue, this study focused on examining the assessment design, practices, and LAL development of EFL assessment experts. In addition, the study explored how contextual and experiential factors have shaped the LAL profiles of the members of this group. By exploring their experiences and practices, the research endeavors to unravel the intricate dynamics of LAL within this specific professional context.

Literature Review

Language Assessment Literacy (LAL): Conceptualizations, models, and frameworks

Perspectives on the conceptualization of LAL have evolved over time, starting with foundational works by Davies (2008), who categorized LAL as comprising skills, knowledge, and principles, and Inbar-Lourie (2008), who emphasized the dimensions of what, how, and why in LAL. A significant advancement came with Fulcher's (2012) comprehensive model, presenting LAL in a hierarchical framework. This framework delineated knowledge, skills, and abilities at the lower level, processes, principles, and concepts at the mid-level, and the

utilization of assessment within broader historical, social, political, and philosophical contexts at the high level.

The pursuit of a definitive LAL framework prompted two significant meta-analyses. Initially, [Xu and Brown \(2016\)](#) conducted a synthesis of 100 studies, culminating in a six-tier conceptual model known as TALiP. This model highlighted both cognitive and emotional dimensions of LAL, particularly by focusing on reshaping the teacher's identity as an assessor. Since this model primarily addressed general assessment literacy among educators, [Gan and Lam \(2022\)](#) undertook a recent examination of the collection of 81 studies that focused solely on LAL. This examination revealed two important shortcomings of LAL studies. The first one was an overemphasis on the language educators' LAL at the cost of neglecting other stakeholders. The second was the formation of LAL benchmarks without accompanying developmental routes.

Another instrumental contribution to LAL conceptualization was made by [Kremmel and Harding \(2020\)](#), who developed a LAL componential framework, leveraging [Taylor's \(2013\)](#) hypothesized LAL components. This empirical and thorough research delineated the following nine separate components of LAL.

1. Crafting and implementing language assessments,
2. Evaluating language pedagogy,
3. Examining assessment policies and local customs,
4. Exploring personal beliefs and attitudes,
5. Employing statistical and research methodologies,
6. Understanding assessment principles and interpretations,
7. Studying language structure and usage,
8. Considering washback effects, and
9. Conducting scoring and rating procedures.

Through exploratory factor analysis, [Kremmel and Harding's \(2020\)](#) study empirically validated its LAL model, providing a foundational basis for its evolving conceptualization. The evolving conceptualization of LAL has continued with [Inbar-Lourie's \(2013\)](#) call for "a multi-dimensional LAL continuum." Among many factors along this continuum, she emphasizes the intricate interplay of various elements in assessment practices, which is also highlighted by [Kremmel and Harding \(2020\)](#). This understanding underscores the interaction of stakeholders' LAL and contextual factors in ensuring effective assessment. Along the same lines, insights from [Villa Larenas & Brunfaut \(2023\)](#) highlight the significance of contextual factors in dealing with the LAL of teacher educators. They conclude that language assessment training and previous experiences with assessment are two influential aspects in developing LAL. This aligns with broader discussions by [Xu and Brown \(2016\)](#) about the intricate interplay between internal cognitive traits, affective beliefs, and external sociocultural influences in the development of AL.

Moreover, Villa Larenas and Brunfaut (2023) examined that various background factors influence teacher educators' language assessment literacy. They made a clear distinction between previous experiences or training—and those currently affecting their practice, like institutional expectations or educational policies. The past context consists of two main elements: first, the training educators have undergone in language assessment—whether through structured courses, workshops, or teacher training programs—and second, their earlier encounters with assessment, particularly in language teaching. The current context, on the other hand, involves the subcategories of macro-sociocultural and micro-institutional factors in the settings where stakeholders work. Xu and Brown (2016) also suggest that what teachers actually do in assessment is often influenced by the need to find a balance between what they believe and what outside forces require such as institutional policies, testing requirements, or curriculum demands. Villa Larenas and Brunfaut (2023) not only agree with this idea but also argue that teacher educators' assessment decisions depend on how they navigate their assessment expertise, their personal understanding of what LAL means, and the specific contexts in which they work. They came to the conclusion that to truly understand language teaching professionals' assessment literacy—and to design effective training programs—it's essential to explore their working environments in depth. This includes looking closely at the local assessment culture, the kind of access they have to assessment-related knowledge, and the beliefs they hold about assessment.

Factors Shaping LAL Development

Research indicates a shift toward a new perspective on teachers' understanding and implementation of assessment that requires a reevaluation of the expectations from them in language assessments (Malone, 2013; Taylor, 2013). These researchers tend to believe that teachers' assessment methods are often shaped by ongoing interactions among various elements, such as the assessment environment, teachers' expertise, beliefs, and past encounters in language teaching, learning, and assessment (Borg, 2003). These complex sociocultural dynamics shape how LAL evolves through its interaction with local culture which, in turn, may impact stakeholders' identities. In addition, personal factors such as individual's awareness of their identity as assessors also play a role in LAL acquisition. However, as Gan and Lam (2022) noted, many studies focus on identifying aspects of LAL that have been developed and improved, while they often overlook the actual process of LAL development. Among the surveyed studies, only four out of 71 delved into the developmental trajectory of LAL within stakeholders' assessment practices, while six studies examined the facilitating or hindering factors affecting LAL development.

Research by scholars like Berry et al. (2019) and Villa Larenas and Brunfaut (2023) has revealed that teacher educators' own experiences with assessment during their education as a student play a role in shaping their attitudes toward language assessment today. This aligns with the findings of Levy-Vered and Alhija (2015) as well as López and Bernal (2009), who found that training in assessment improves how teachers think about assessment and in turn, strengthens their assessment literacy.

Other investigations also suggest that training influences teachers' assessment practice and that language teachers are more likely to use the assessment procedures for which they have

received training (Reynolds-Keefers, 2010; Rohl, 1999). As an example, Xu and Brown (2016) argued that conceptions of assessment filter the knowledge acquired through training and stated, “teachers tend to adopt new knowledge, ideas, and strategies of assessment that are congruent with their conceptions of assessment while rejecting those that are not” (p. 156).

Giraldo (2021) conducted a comprehensive review of the literature on LAL and found that data from all 14 studies surveyed suggest that teachers have received training in various educational settings and languages. Across different teaching contexts and languages, the research points to the importance of the need to strengthen teachers’ language assessment literacy especially by focusing on the skills aspect through applying reflective strategies to test design and creation, making this a main element of teacher training and development. Crusan et al. (2016) found that teachers’ self-perceived language assessment literacy depends heavily on personal experiences, such as their linguistic background and teaching experience. Building on these findings, Yan et al. (2018) encouraged language teachers to consider their own LAL development and discovered that most teachers were more engaged with hand-on experiences of assessment than with its theoretical underpinnings. Along with contextual and background factors, improving assessment literacy may also depend on self-reflection and guided learning, both of which are core aspects of professional growth. Scarino (2013) pointed out that when language teachers regularly reflect on their assessment practices, it brings their underlying beliefs about assessment to light. This kind of reflection can help them reflect on the role of assessment within teaching-learning process. Beyond self-reflection, apprenticeship also plays a key role in professional development. Beginners in the field, in particular, stand to learn from valuable insights and skills of experienced colleagues (Crandall, 2000; Johnson, 2006). Apart from contextual factors, teachers’ LAL development is also affected by their background in assessment, professional development, and classroom practice. Experts in the fields of language assessment and teacher education assert that reflecting on assessment-related experiences can enhance teachers’ professional agency and facilitate their shift from instructors to professional teacher-assessors (Crusan et al., 2016; Lee, 2010).

Scarino (2013) and Vogta and Tsagari (2014) emphasized that teachers gain assessment competence via practical experience, progressively forming assessment instincts understandings of assessment and acquiring clearer insights into their professional development needs (p. 382). Babaii and Asadnia (2019) brought attention to the significance of reflective practices in strengthening EFL teachers’ language assessment literacy (LAL), emphasizing the need for teachers to engage more earnestly with their roles as language assessors. It promotes sustained collaboration between university educators and institute-based trainers, facilitate the integration of research findings, and engage in hands-on practice. With regard to teaching experience, Crusan et al. (2016) reported that teachers with different lengths of experience reported varying levels of confidence in their assessment knowledge and practices. Notably, among the teachers surveyed, less experienced teachers expressed higher levels of confidence in understanding and application of assessment when compared with their more experienced peers. They identify several reasons to explain the variation: (1) teacher preparation program has increasingly expanded assessment training over time; (2) ongoing assessment training for practicing teachers is often insufficient; and (3) experienced teachers may be less receptive to new assessment training compared to their beginner counterparts.

Additionally, earlier research has shown that teachers' assessment practices are shaped by their classroom setting, institutional policies, education mandates, and the broader sociocultural beliefs about language instruction and testing. For example, Ruecker et al. (2014) observed that teachers working in K-12 and EFL settings often use summative grammar exams to assess second language writing performance. Conversely, teachers in postsecondary ESL contexts tend to use open-ended writing assignments and alternative assessment methods, such as reflective writing, blogs, emails, and discussion boards.

Even though the study did not delve into the causes of these differences in depth, it is reasonable to suggest that factors such as learner characteristics, teachers' workload, and the availability of assessment resources may explain these variations. With respect to institutional and policy demands, Rea-Dickins (2001) studied L2 teachers' classroom assessments and pinpointed three key functions of how teachers construct assessment.

Rea-Dickins (2001) identified three core functions of institutional assessment, each corresponding to distinct stakeholder groups: (1) Pedagogical function (for learners): detecting learner requirements and fostering language development; (2) Instructional function (for teachers): tracking learner progress and determining teaching strategies; and (3) Administrative function (for institutional authorities): evaluating learning outcomes in line with policy benchmarks. The role of educational policies and sociocultural values was noted in a comprehensive study examining foreign language instructors' language assessment literacy (LAL) and professional training requirements across six European countries (Vogta & Tsagari, 2014). While teachers expressed keen interest in various components of classroom-based assessment practices, international differences emerged regarding their priorities. For instance, teachers in Greece prioritized formal assessment techniques, whereas, teachers in Germany placed greater emphasis on reading and writing assessment, as these skills were main emphases to high stakes national examinations.

Additionally, fostering collaboration among diverse assessment practitioners is essential for advancing LAL (Lee et al., 2021). Through collaboration, language educators engage in teamwork with peers, students, assessment trainers, and teacher educators. This collaborative approach creates a social setting, where stakeholders interact to develop a shared understanding of LAL acquisition. Building on this premiss, Crusan et al. (2016) categorised factors influencing teachers' LAL development into two main groups: first, contextual factors that are related to the broader educational, social, cultural, political, and historical aspects shaping the assessment culture within a specific context. Second, experiential factors are linked to individual teachers' backgrounds, training, and experience in language teaching and assessment. By employing such a framework, one can identify influential factors within these categories for a deeper understanding of the LAL concept and its developmental pathways for EFL assessment professionals.

In essence, the literature review accentuates the need for an elaborate research framework that encompasses different stakeholders including EFL assessment experts within the broader landscape of LAL research. The subsequent section outlines the methodological approach adopted to delve into and depict the LAL of this specific professional group, aiming to bridge

the existing gap in understanding the unique contributions of EFL assessment experts to L2 assessment. To this end, three research questions were addressed in this study:

1. How do EFL assessment experts conceptualize and define language assessment literacy (LAL)?
2. How have they developed their language assessment literacy (LAL)?
3. What are the primary factors perceived by experts that influence the design and implementation of EFL classroom assessments?

Method

Conceptual framework of the study

To achieve the study's objectives, several conceptual frameworks were employed to identify as many factors impacting LAL as possible. Drawing upon Taylor's (2013) hypothesized LAL components, a significant advancement was achieved by Kremmel and Harding (2020), who developed a LAL componential framework. This study, distinguished for its empirical foundation and comprehensive scope, identified nine distinct elements of LAL. These factors included creating and conducting language assessments, assessment integration in instruction, policy frameworks, and contextual approaches, individual perspective, research methods and data analysis methods, assessment principles and interpretation, language structure and use, washback, and marking procedures. Villa Larenas and Brunfaut (2023) effectively included these elements into their conceptual framework, highlighting the differences between past influences and those emerging from in the present context. The historical context examines two key aspects closely: first, the structured training that stakeholders acquired in language assessment through educational or professional programs, and second, their practical experiences with assessment. In contrast, Xu and Brown (2016) outlined that the present context integrates both broader sociocultural contexts and more localized, institutional policies.

Participants and Setting

The study involved 12 experienced both male and female EFL teachers who were teaching at universities across Iran. The participants with an age range of 45 to 65 had more than twenty years of experience in teaching English. Following theoretical sampling to recruit the informants, the researchers could reach saturation after interviewing 12 assessment experts. The selection criteria for identifying participants as experts in this study were based on the proficiency model established by Chi (2006). For this study, language assessment experts are professionals whose main responsibilities involve a range of language assessment tasks including administering language tests, teaching testing courses at the undergraduate or graduate levels, conducting academic research in language testing, and providing training in language assessment. The expert participants enjoyed higher levels of expertise since they were all instructors affiliated with Iranian universities, specializing specifically in language testing research. Besides, they actively participate in developing and validating nationwide tests as part of their professional obligations.

Following the guidance of Bogdan and Biklen (2007), only those who agreed to participate voluntarily were included in the study, after providing a clear overview about how the study

would be conducted. Participants were provided with a brief and clear explanation of the study's purpose, methodology, and how the collected data would be collected. Further, they were informed that they could withdraw from participation at any time and for any reason. They could also skip the questions they would not feel comfortable with and leave them unanswered. A significant concern for a qualitative researcher is to address the ethical issues involved in the research process (Ary et al., 2013). Attempts were made to observe all ethical considerations (consent, anonymity, and confidentiality). The participants' demographic details are summarized in Table 1.

Table 1. *Demographic Information of the Participants*

Number (percent)		
Education	Ph. D	12 (100)
Gender	Male	5 (41.66%)
	Female	7 (58.33%)
Experience	20-23 years	3 (24.99%)
	23-26 years	4 (33.33%)
	Above 26	5 (41.66.6%)

In the Iranian educational context, students who choose to major in teaching English as a foreign language (TEFL) are required to take courses in language testing. This pedagogical requirement persists throughout their academic trajectory, with testing courses extending into the Master's and Doctoral studies. Upon graduation, these individuals are entrusted with the responsibility of crafting assessments for different language learning environments. These experts are called upon to design and evaluate language tests for specific purposes, including university admissions, selection tests for higher education institutions, and specialized English for Specific Purposes (ESP) courses.

Data Collection Procedure

Data were obtained from three distinct sources: analysis of the syllabi, narrative framing, and cognitive interviews. One of the primary data collection methods involved collecting syllabi prepared by language assessment experts for their assessment courses. Each expert provided two syllabi for analysis. Their content underwent meticulous examination to pinpoint the experts' perceptions of various language assessment components that they incorporated into their syllabi. We could also make inferences about potential reasons for the inclusion of the components, the way expert utilized them, and the rationale behind their usage. We also situated narrative frames within the context of instructors' LAL and practice that they developed through training and experiences. The employment of narratives originated the idea that when experts share their experiences, they engage in reflection and understanding of those experiences; thereby, they deepen their grasp of their teaching knowledge and practice (Barkhuizen & Wette, 2008, 374). This study adopted narrative frames as they offer structured guidance and support for form and content (Barkhuizen & Wette, 2008, 376). To examine the content validity of the narrative frames, a panel of three experts in the relevant field reviewed the frames for final refinements. Then, a pilot study was conducted with a group of three

participants with characteristics similar to those of the target population to identify and address any ambiguities. This approach provided the participants with sufficient time to recall and reflect on various aspects related to assessment beliefs, knowledge, training, experiences, and perceptions of the factors influencing the design and practice of language assessment.

To delve deeper into the examination of the syllabi and experts' narratives, cognitive interviews were conducted. Cognitive interviews are rooted in "psychological principles of remembering and retrieving information from memory" (Stein & Memon, 2006, 597), enhancing the depth of understanding and data reliability (Stein & Memon, 2006). During the interview phase, the researchers posed questions about the content of the syllabi and the narratives to get insights into how the experts addressed different aspects of LAL. These syllabi served as conversation prompts, facilitating discussions on topics related to the participants' language assessment practices. During the conversations, the participants delved into discussions about the purpose, development, uses, and perceived quality of their syllabi. Additionally, discussions often extended to significant issues concerning institutional factors, assessment culture, and the language assessment expert community. These face-to-face interviews had an average duration of 20 minutes per participant. The components proposed by Villa Larenas and Brunfaut (2023) served as the foundation for designing the semi-structured interview guide, providing a structured framework for inquiry while also allowing flexibility for interviewees to elaborate on topics they considered pertinent. The interview questions were validated by three experts in the field of applied linguistics, who suggested minor revisions before its implementation.

Data Analysis

For the qualitative phase, we used Interpretative Phenomenological Analysis (IPA) to gain in-depth understanding of the experts' Language Assessment Literacy (LAL), previous experiences, and training, focusing on their assessment knowledge beliefs, and practices. The analysis encompassed the language assessment syllabi prepared by the experts, the framed narratives, and subsequent interviews. These data were subjected to analysis using specific coding criteria, as detailed in the section concerning interviews. The researchers begin by listening to each interview once before transcribing it during a second, more detailed review. Transcribing the interviews took an average of 50 minutes each, and the follow-up discussions varied between 50 and 90 minutes. To do the analysis systematically, the researcher followed the steps recommended by Gao and Zhang (2020), which are: (1) checking the data for any instance of ambiguity, irrelevance, or error, and modifying them; (2) generating open codes from the data, (3) generating subthemes/axial codes from the data, (4) grouping subthemes under more general themes/selective codes, and (5) writing a detailed and finalized report of the analysis process. The inter-coder agreement was checked with 20% of the codes that were examined by an outsider, who was an expert in applied linguistics. Accordingly, in this study, a total of 145 codes were obtained. Twenty-nine codes were coded by the outsider. The two coders disagreed on one code only. The inter-coder agreement coefficient obtained was 97%. The two coders discussed the points of disagreement and finalized the code system. MAXMAP properties in MAXQDA were adopted to create figures for presenting the themes, categories,

and codes. Moreover, frequency counts were calculated to show the occurrences of codes (Table 2).

Table 2. *Conceptualization of Language Assessment Literacy by EFL Assessment Experts*

Components of LAL identified.	Code Frequency
Language assessment knowledge	12
Contextual variables	8
Conceptions of language assessment	5
Language assessment practices	12
Language assessment learning	10
Reflectivity	12
Agency	8
Total	67

Findings

EFL Assessment Experts' Conceptualization of LAL

As displayed in Table 2, the participants as experts in the domain of language assessment, collectively highlighted the multifaceted factors influencing LAL in line with frameworks proposed by Villa Larenas & Brunfaut (2023) and Kremmel and Harding (2020). These frameworks systematically depict the fundamental constituents of LAL. The emergent themes derived from qualitative data analysis included almost all constituents delineated within these frameworks. The study included several key areas such as language assessment knowledge, contextual factors, views on language testing, practical procedures and methods, and the process of learning about language assessment. Some of the comments made by the participants demonstrate the significance of these aspects of LAL.

"By having the theoretical knowledge, I try to utilize assessment techniques that are applicable in my classes."

"Assessment should also be free from bias."

"Of course, a test should be both valid and reliable"

"I would incorporate both communicative and highly grammatical tests where both meticulous knowledge of structure as well as communication could be tested."

"The one that would be at the service of learning and teaching."

"Ideal assessment or test should consider the purpose of the test and its functions."

"Assessment partly depends on institute rules and regulations."

During the data analysis, two additional thematic dimensions surfaced, which had not been explicitly encapsulated within prevailing theoretical frameworks. These newly emergent themes received high attention from the cohort of eight participants, who collectively underscored the pivotal role ascribed to the agency and reflectivity of language assessors. As one of the participants put it:

"As an active member of this community, we must actively utilize our agency to navigate the complexities of our practice and make meaningful progress in assessment methodologies."

Another participant articulated that:

“We need to be reflective enough as it leads to continuous improvement in language assessment practices by fostering critical self-awareness and informed decision-making”.

Participants also highlighted the important role of self-study and Teacher Training Courses (TTC) in developing LAL, echoing responses from experts regarding the study's second research question (see Table 2). This proactive engagement signifies a notable level of professional agency among experts, demonstrating their active involvement in shaping language assessment practices. In scholarly discourse, professional agency is recognized as the capacity for intentional action and transformative practice within one's professional domain (Bandura, 2001). Moreover; According to Bandura (2001), self-reflectiveness about one's capabilities contributes to the development of human agency. Through self-study, experts engage in reflective inquiry and critical examination of assessment theories, methodologies, and practices, thereby fostering a deepened understanding of the complexities inherent in language assessment.

In short, in addition to the components of LAL identified in the existing models of LAL, the frequent discourse on the concepts of reflectivity and agency among the experts seem to be significant in shaping the content of LAL. These two intertwined notions are recurrent themes within the experts' discussions, reflecting their profound impact on the development and refinement of language assessment practices.

Development of Language Assessment Literacy (LAL) among EFL Assessment Experts

Regarding the second research question, according to Table 3 and Figure 2 below, the experts received training on assessment in three ways “Passing assessment courses in university (f = 10), Increasing knowledge through self-study (f = 5), and Attending TTC (f = 3)”. The influence of previous training and experiences on experts' current assessment practices was evident in their responses. The richness of their educational background including university courses, self-directed study, and TTC, was cited as shaping their assessment methodologies. This affirmation aligns with the conclusions drawn by Levy-Vered and Alhija (2015) as well as López and Bernal (2009), indicating that assessment training has a profound influence on teachers' assessment perceptions. The experts underscored the significance of self-directed learning in influencing their current assessment practices. Five participants highlighted the impact of independent study, emphasizing how continuous efforts to stay informed about the latest developments in assessment through self-study contributed to their evolving approaches. In this respect, some experts for example mentioned that:

“I keep up with the latest assessment developments, reading assessment journals, and assessment textbooks.”

“I have just passed language assessment courses in BA, MA, PhD courses. Not extra assessment training.”

“Not all that were presented to us through textbooks could be applied to practice, so I conduct informal needs assessments to know what my students require so that I will tailor instruction to assessment.”

"I have been influenced by my experience as a student. But I have tried to let students to evaluate their peers or themselves on their knowledge, skills or the strategy types that they use for learning."

"Definitely my own experience has great influence on my own style of assessing students, in addition students are also accustomed to the traditional assessment and prefer to be assessed in the same way. In fact, I am not familiar to other ways of assessing and I am continuing the same trends."

"I have taken some teacher training courses on it."

Additionally, the experts reported their previous experience of nine categories of assessment as a student, which are Summative tests (F = 12), Written tests (F = 8), Oral tests (F = 5), Formative assessment (F = 3), Traditional assessment (F = 2), Achievement tests (F = 1), Process-oriented assessment (F = 1), Grade-based assessment (F = 1), and Translation tests (F = 1). More specifically, for instance, they mentioned that:

"End-of-course exams were very common."

"We were assessed through a written examination (which focused on our general English knowledge) once a month."

"The oral interviews which focused on our teaching abilities and skills were conducted once a week."

"We were assessed at different intervals by the instructor."

"The teacher gave a mark to our performance."

Table 3. Development of Language Assessment Literacy among experts

Previous training categories	Code Frequency
Passing assessment courses at the university	10
Increasing knowledge through self-study	5
Attending TTC	3
Total	18
Previous experiences as a student category	Code Frequency
Summative tests	10
Written tests	8
Oral tests	5
Formative assessment	3
Traditional assessment	2
Achievement tests	1
Process-oriented assessment	1
Grade-based assessment	1
Translation tests	1
Total	34

Figure 2 presents a visual representation of the frequencies related to experts' previous training and experiences as students, shedding light on the diverse pathways through which these professionals have developed their assessment knowledge.

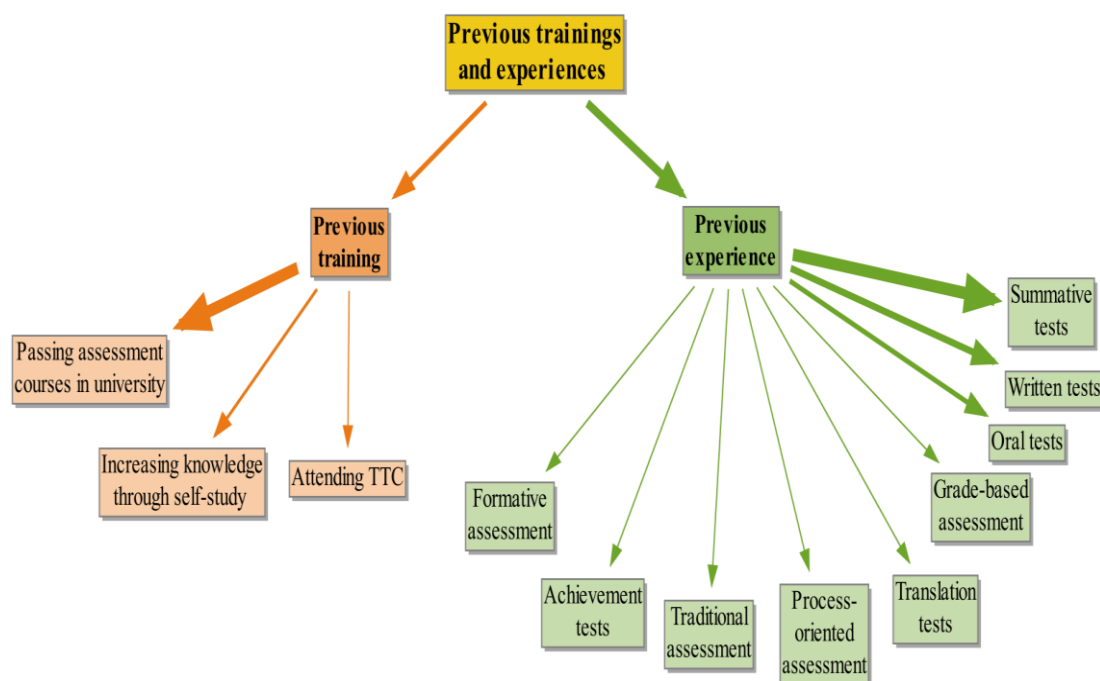


Figure 1. Development of Language Assessment Literacy among experts

The main factors influencing the design and practice of EFL classroom assessments

From experts' perspectives, 11 factors influence design and practice of their EFL classroom assessments, which are Motivation ($F = 6$), Workload ($F = 3$) Teacher training courses ($F = 3$), Access to teaching and learning resources ($F = 3$), Personal interest ($F = 2$), Student needs ($F = 1$), Teacher decisions ($F = 1$), Rules of institution ($F = 1$), Organizational factors ($F = 1$), Lack of discipline ($F = 1$), and Student laziness ($F = 1$). In this regard, the experts mentioned that:

"Lack of discipline impacts my assessment."

"Assessment partly depends on institute rules and regulations."

"Assessment depends to some extent on my own choice."

"I need to attend to students' needs as well."

"Student laziness influences what I do."

"The personal interest along with access to learning and teaching resources, opportunities for professional development workshops, workload issues, and student motivation can affect the assessment techniques I use."

As a distinctive aspect of the current investigation, our endeavor involved elucidating the hierarchical structure of factors influencing the substance of LAL among experts. These findings, presented in Table 3 and Figure 3, are supplemented by the frequency distribution of these factors.

Table 3. *The factors influencing the design and practice of experts' EFL classroom assessments*

Factors influencing design and practice of classroom assessments	Code Frequency
Motivation	6
Workload	3
Teacher training courses	3
Access to teaching and learning resources	3
Personal interest	2
Student needs	1
Teacher decisions	1
Rules of institution	1
Organizational factors	1
Lack of discipline	1
Student laziness	1
Total	23

Figure 3 illustrates the varied factors perceived by experts as influential in shaping the design and practice of EFL classroom assessments. These insights provide a comprehensive view of the multifaceted considerations that impact the assessment processes.

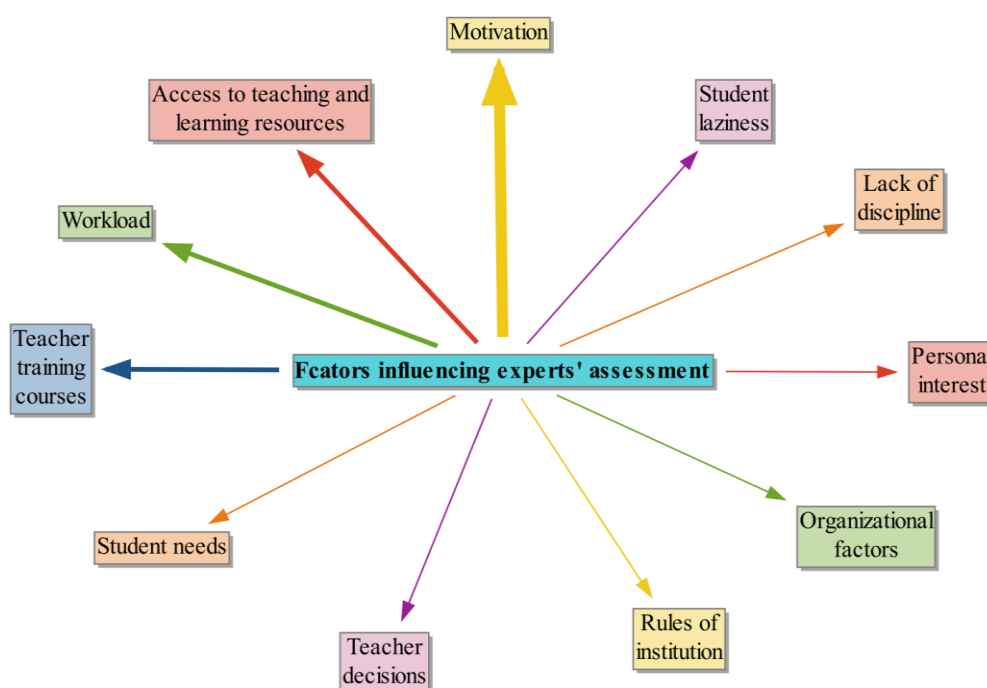
**Figure 2.** *The factors influencing the design and practice of experts' EFL classroom assessments*

Table 3 and Figure 3 reveal a diverse array of factors influencing the design and practice of experts' EFL classroom assessments. These factors are both internal such as educators' personal interest and motivation, and external including workload, access to resources, and institutional rules. Importantly, educators' participation in teacher training courses equips them with the necessary skills for effective assessment practices, highlighting the importance of ongoing professional development in enhancing assessment quality and promoting student learning.

Discussion

The dearth of studies focused on EFL assessment experts is a notable gap in the literature, despite broader discussions on language assessment practices. Gan and Lam (2022) highlight the underrepresentation of various stakeholders in LAL studies, including EFL assessment experts. Collaborative efforts among assessment-related practitioners are essential for the cultivation of LAL (Lee et al., 2021), fostering mutual learning and the emergence of new assessment processes. Building on Villa Larenas & Brunfaut's (2023) as well as Kremmel and Harding's (2020) framework, we explored influential factors in both contextual and experiential domains to enhance our understanding of LAL development paths for EFL assessment experts. Through qualitative data analysis, emergent themes encompassing key constituents of LAL were identified. The findings were consistent with those of prior studies (Gan & Lam, 2022; Inbar-Lourie, 2013; Kremmel & Harding, 2020; Villa Larenas & Brunfaut, 2023; Xu & Brown, 2016).

Notably, participants highlighted the pivotal role of agency and reflectivity in language assessment practices, two dimensions not explicitly captured by existing frameworks. Miller and Gkonou (2018) assert that engagement in assessment practices shapes teachers' agency and contributes to their understanding of how assessment should be conducted. For example, Gu (2013) discovered that the LAL of an experienced Chinese EFL teacher is constrained by prescribed curriculum standards, limiting their autonomy and agency. Therefore, to improve EFL teachers' LAL, as Berry et al. (2019) suggest, incorporating both experiential and reflective elements into LAL training seems inevitable. In their investigation, Tian et al. (2021) employed on-task dialogues and demonstrated that as participants progressed through reflective dialogues, they gained confidence in their LAL, and assumed greater authority in addressing assessment challenges. Their study also showed that participants became increasingly mindful of adopting a perspective on language assessment that allows for tailored alternative assessments to meet pedagogical requirements. They recommend that frontline educators should be afforded opportunities to cultivate their LAL by engaging with fellow professionals in reflective practices within and outside their cultural spheres.

The study also found that the way experts conduct assessment is strongly influenced by their previous training and professional background. This supports Villa Larenas and Brunfaut's (2023) framework in which they divided the contextual factors affecting teacher educators' Language Assessment Literacy (LAL) into past and present contexts. The past context refers to the language assessment training participants have completed such as formal courses, workshops, or professional development as well as their earlier hands-on experience with assessment. The present context encompasses wide-ranging social and cultural influences as well as localized institutional factors that stem from the socio-political and educational environments where stakeholders operate.

The participants' inclination towards formal university courses and self-directed learning aligns with the conclusions drawn in different studies (Lam 2015; Jin 2010; Villa Larenas & Brunfaut 2023). The study by Kremmel and Harding (2020) emphasizes the importance of self-directed learning and continuous professional development in shaping educators' assessment practices. This finding aligns with Villa Larenas and Brunfaut's (2023) observation that teacher

educators' previous experiences with assessment as students shape their attitudes toward language assessment. Furthermore, the nuances of experts' experiences as students unfolded in various dimensions, including oral tests, formative assessment, and traditional assessment methods.

These findings align with those of previous research emphasizing the intricate and multifaceted nature of factors influencing educators' assessment design (Coombe et al., 2020; Gan & Lam 2022; Giraldo, 2021; Vogt, Tsagari, & Spanoudis 2020). Participants in our study placed significant emphasis on teacher motivation, a factor that corresponds with the framework developed by Villa Larenas & Brunfaut (2023) under the category of emotions and beliefs. The interplay of contextual, experiential, and personal factors in LAL development is highlighted, emphasizing the complex sociocultural dynamics that shape stakeholders' identities concerning assessment practices.

Conclusion

By providing nuanced insights into the knowledge, conceptions, and practices of EFL assessment experts, this study contributes to the broader discourse on language assessment practices and literacy. The findings of our study offer a reasonable understanding of LAL among Iranian EFL assessment experts, enriching the existing literature on this topic. As discussed by Pill and Harding (2013) and Taylor (2013), our study also confirms the notion that different stakeholders within the language assessment community possess varying levels of LAL, necessitating tailored approaches to LAL development. By examining LAL among assessment experts, our study contributes to the empirical validation and refinement of LAL frameworks proposed by previous scholars. This empirical grounding substantiates the theoretical foundations of LAL and provides valuable insights for the design of targeted LAL training initiatives. Moving forward, future research endeavors should aim to further explore the dynamic nature of LAL and its implications for language education policy and practice. Furthermore, comparative studies across different educational contexts and cultural settings could enrich our understanding of the LAL components in diverse contexts.

References

- Abrar-ul-Hassan, S., & Nassaji, H. (2024). Rescoping language assessment literacy: An expanded perspective. *System*, 120, 103195.
- Ary, D., Jacobs, L. C., Sorensen, C., & Razavieh, A. (2010). *Introduction to Research in Education*. Cengage Learning.
- Babaii, E., & Asadnia, F. (2019). A long walk to language assessment literacy: EFL teachers' reflection on language assessment research and practice. *Reflective Practice*, 20(6), 745–760. <https://doi.org/10.1080/14623943.2019.1688779>
- Baker, B. A., & Riches, C. (2018). The development of EFL examinations in Haiti: Collaboration and language assessment literacy development. *Language Testing*, 35(4), 557–581. <https://doi.org/10.1177/0265532217716732>
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52(1), 1–26.
- Barkhuizen, G., & Wette, R. (2008). Narrative frames for investigating the experiences of language teachers. *System*, 36(3), 372–387. <https://doi.org/10.1016/j.system.2008.02.002>.
- Berry, V., Sheehan, S., & Munro, S. (2019). What does language assessment literacy mean to teachers? *ELT Journal*, 73(2), 113–123. <https://doi.org/10.1093/elt/ccy055>
- Bogdan, R.C. and Biklen, S.K. (2007) *Qualitative Research for Education: An Introduction to Theory and Methods*. Allyn & Bacon, Boston.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(02), 81–109. <https://doi.org/10.1017/S0261444803001903>
- Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21–30). Cambridge University Press.
- Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: What do we need to learn, unlearn, and relearn? *Language Testing in Asia*, 10. Article 3. <https://doi.org/10.1186/s40468-020-00101-6>
- Crandall, J. J. (2000). Language teacher education. *Annual Review of Applied Linguistics*, 20, 34–55. <https://doi.org/10.1017/S0267190500200032>
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. 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>
- Dastgoshadeh, A., & Jalilzadeh, K. (2024, In press). Unraveling the nexus between teachers' past assessment experiences, training, and current assessment practices: A case of Iranian EFL teachers. *The Journal of Asia TEFL*.
- Davies, A. (2008). *Textbook trends in teaching language testing*. *Language Testing*, 25(3), 327–347. <https://doi.org/10.1177/0265532208090156>
- Eraut, M. (2004). Transfer of knowledge between education and workplace settings. In H. Rainbird, A. Fuller, & H. Munro (Eds.), *Workplace learning in context* (pp. 201–221). Routledge.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
- Gan, L., & Lam, R. (2022). A review on language assessment literacy: Trends, foci, and contributions. *Language Assessment Quarterly*, 19(5), 503–525. <https://doi.org/10.1080/15434303.2022.2128802>

- Gao, L. X., & Zhang, L. J. (2020). Teacher learning in difficult times: Examining foreign language teachers' cognitions about online teaching to tide over COVID-19. *Frontiers in Psychology*, 11, 2396. <https://doi.org/10.3389/fpsyg.2020.549653>
- Giraldo, F. (2021). A reflection on initiatives for teachers' professional development through language assessment literacy. *Profile: Issues in Teachers' Professional Development*, 23(1), 197–213. <https://doi.org/10.15446/profile.v23n1.83094>
- Gu, P. Y. (2013). The unbearable lightness of the curriculum: What drives the assessment practices of a teacher of English as a foreign language in a Chinese secondary school? *Assessment in Education: Principles, Policy & Practice*, 21(3), 286–305. <https://doi.org/10.1080/0969594x.2013.836076>
- Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 413–428). Mouton de Gruyter.
- Hasselgreen, A., Carlsen, C., & Helness, H. (2004). *European survey of language testing and assessment needs. Part one: General findings*. <http://www.ealta.eu.org/resources.htm>
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402. <https://doi.org/10.1177/0265532208090158>
- Inbar-Lourie, O. (2013). Guest editorial to the special issue on language assessment literacy. *Language Testing*, 30(3), 301–307. <https://doi.org/10.1177/0265532213480126>
- Inbar-Lourie, O. (2017). Language assessment literacy. In E. Shohamy, I. Or, & S. May (Eds.), *Language testing and assessment, Encyclopedia of language and education* (pp. 257–286). Springer International Publishing. https://doi.org/10.1007/978-3-319-02261-1_19
- Jin, Y. (2010). The place of LTA in the professional preparation of foreign language teachers in China. *Language Testing*, 27(4), 555–584. <https://doi.org/10.1177/0265532209351431>
- Johnson, K. E. (2006). The sociocultural turn and its challenges for second language teacher education. *TESOL Quarterly*, 40(1), 235–257. <https://doi.org/10.2307/40264518>
- 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(1), 100–120. <https://doi.org/10.1080/15434303.2019.1674855>
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169–197. <https://doi.org/10.1177/0265532214554321>
- Lee, I. (2010). Writing teacher education and teacher learning: Testimonies of four EFL teachers. *Journal of Second Language Writing*, 19(3), 143–157.
- Lee, J., Butler, Y. G., & Peng, X. (2021). Multiple stakeholder interaction to enhance pre-service teachers' language assessment literacy. *Languages*, 6(4), 213. <https://doi.org/10.3390/languages6040213>
- Levy-Vered, A., & Alhija, F. (2015). Modeling beginning teachers' assessment literacy: The contribution of training, self-efficacy, and conceptions of assessment. *Educational Research and Evaluation*, 21, 378–406. <https://doi.org/10.1080/13803611.2015.1117980>
- López-Mendoza, A. A., & Bernal-Arandia, R. (2009). Language testing in Colombia: A call for more teacher education and teacher training in language assessment. *Profile: Issues in Teacher's Professional Development*, 11(2), 55–70.
- Malone, M. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329–344. <https://doi.org/10.1177/0265532213480129>
- Miller, E. R., & Gkonou, C. (2018). Language teacher agency, emotion labor and emotional rewards in tertiary-level English language programmes. *System*, 79, 49–59. <https://doi.org/10.1016/j.system.2018.03.002>

- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381–402. <https://doi.org/10.1177/0265532213480337>
- Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, 18(4), 429–462.
- Reynolds-Keefers, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment, Research and Evaluation*, 15(8), 1–9. Retrieved from <http://pareonline.net/pdf/v15n8.pdf>
- Rohl, M. (1999). Profiling ESL children: How teachers interpret and use national and state assessment frameworks. *Queensland Journal of Educational Research*, 15(1), 113–122. Retrieved from <http://www.iier.org.au/qjer/qjer15/rohl.html>
- Ruecker, T., Shapiro, S., Johnson, E. N., & Tardy, C. M. (2014). Exploring the linguistic and institutional contexts of writing instruction in TESOL. *TESOL Quarterly*, 48(2), 400–412.
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and teacher learning. *Language Testing*, 30(3), 309–327. <https://doi.org/10.1177/0265532213480128>
- Stein, L. M., & Memon, A. (2006). Testing the efficacy of the cognitive interview in a developing country. *Applied Cognitive Psychology*, 20, 597–605. <https://doi.org/10.1002/acp.1211>.
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21–36. <https://doi.org/10.1017/S0267190509090035>
- 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/10.1177/0265532213480338s>
- Tian, W., Louw, S., & Khan, M. K. (2021). COVID-19 as a critical incident: Reflection on language assessment literacy and the need for radical changes. *System*, 103, Article 102682.
- Tsagari, D., & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges, and future prospects. *Papers in Language Testing and Assessment*, 6(1), 41–63.
- Villa Larenas, S., & Brunfaut, T. (2023). But who trains the language teacher educator who trains the language teacher? An empirical investigation of Chilean EFL teacher educators' language assessment literacy. *Language Testing*, 40(3), 463–492.
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374–402. <http://doi.org/10.1080/15434303.2014.960046>
- Vogt, K., Tsagari, D., & Spanoudis, G. (2020). What do teachers think they want? A comparative study of in-service language teachers' beliefs on LAL training needs. *Language Assessment Quarterly*, 17(4), 386–409. <https://doi.org/10.1080/15434303.2020.1781128>
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualisation. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Yan, X., & Fan, J. (2021). “Am I qualified to be a language tester? Understanding the development of language assessment literacy across three stakeholder groups: Understanding the development of language assessment literacy across three stakeholder groups. *Language Testing*, 38(2), 219–246. <https://doi.org/10.1177/0265532220929924>
- Yan, X., Zhang, C., & Fan, J. S. (2018). “Assessment knowledge is important, but ...” How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System*, 74, 158–168. <https://doi.org/10.1016/j.system.2018.03.003>