

Teachers' Conceptions of Language Assessment: Affective and Theoretical Knowledge Dimensions of Language Assessment Literacy Model

Ebrahim Mohammadkhah¹, Gholam Reza Kiany^{2*}, Zia Tajeddin³,
Parvaneh ShayesteFar⁴

Received: December 2021

Accepted: March 2022

Abstract

The contemporary era of learning-oriented assessment (LOA) demands teacher professional efforts to appropriately and accurately assess learners' attainment and use the assessment results for the enhancement of learning. In second/foreign language (L2) discipline, this has recently brought language assessment literacy (LAL) to the forefront, emphasizing what assessment competencies L2 teachers with special needs, such as English as a Foreign Language (EFL) teachers, require to possess to be assessment literate. Given this, teachers' mastery of assessment practical and technical skills, as prescribed by the exiting LAL models, could not be entirely complete nor enhanced without their conceptions of assessment knowledge base and attitudes or beliefs about language assessment system. In light of this importance, the development of a LAL framework that helps better measure EFL teachers' affective and theoretical/conceptual dimensions of assessment is required. To illuminate these dimensions, the present study took insight from Xu and Brown's (2016) LAL model into the development and validation of a LAL scale which could have relevance to Iranian EFL contexts. To provide empirical support for the utility and validity of the scale, both exploratory and confirmatory factor analyses (EFA and CFA with SPSS and AMOS) were performed on data obtained from a sample of 213 Iranian EFL teachers. The results showed satisfactory internal consistency and construct validity of the scale, suggesting that it has the potential to be used for assessing EFL teachers' LAL conceptions and levels. The paper concludes with the importance of teachers' conceptions of LAL dimensions, and finally presents implications of the findings.

Keywords: Affective and Theoretical Dimensions; Language Assessment Literacy; LAL Scale; Teacher Assessment Knowledge Base

¹ Department of English Language Teaching, Faculty of Humanities, Tarbiat Modares University, Tehran, Iran. E-mail: muhammadkhah.i@gmail.com

^{*2} Department of English Language Teaching, Faculty of Humanities, Tarbiat Modares University, Tehran, Iran.

Corresponding Author, E-mail: kiany_gh@modares.ac.ir

³ Department of English Language Teaching, Faculty of Humanities, Tarbiat Modares University, Tehran, Iran. E-mail: tajeddinz@modares.ac.ir

⁴ Department of English Language Teaching, Farhangian Teacher Education University, Tehran, Iran. E-mail: parishayeste@cfu.ac.ir

1. Introduction

In the current conceptualization of assessment as 'assessment for learning' or learning-oriented assessment (LOA), teachers are expected to possess an adequate level of theoretical, practical or contextual, and epistemological knowledge about assessment (Fulcher, 2021; Hamp-Lyons, 2017). Upon this, the successful implementation of LOA in language classrooms requires language teachers to bridge a range of assessment theoretical and practical skills they have at their disposal to "harness assessment to enhance the learning experience" (Fulcher, 2021, p. 34). This clearly calls for attention to teachers' language assessment literacy (LAL) as "one of the cornerstones needed for successful implementation of educational policy" (Gebriel, 2021, p. 116) within a learning-oriented setting. Nonetheless, the current realities and priorities of LOA implementation still reveal more focus on psychometric- or measurement-based usage of assessment (Carless, 2015) and assessment technical skills as part of knowledge base dimension of LAL than epistemological, affective, and philosophical approaches as another important dimension. The important role of teachers' LAL in assessment and decision-making processes demands a LAL view which accommodates not only knowledge base (i.e., theoretical and technical competencies) but also conceptions of assessment (including epistemological beliefs, and affective components). Conceptions of assessment, as argued by Barnes, et al. (2015), filter and interpret teachers' assessment knowledge base, and mediate their theoretical knowledge base competency and its implementation (Xu & Brown, 2016). Although teacher conceptions of LAL have been acknowledged by some studies, few works have included them as the main dimension of their LAL models.

There is evidence that assessment of language learning has far-reaching consequences for the students and other stakeholders (e.g., parents, policymakers, and teachers) involved in language learning programs, especially when high-stakes decisions are made based on the assessment results. Therefore, teachers' knowledge of assessment, as an important part of their LAL, is central to their assessment practices and the related decisions they make on the basis of assessment data. However, in LOA perspective, teachers' mastery of a wide range of assessment skills and crafts could not be entirely enhanced nor complete without their clear conceptions and conceptualization of assessment approaches and disciplinary knowledge, and epistemological beliefs about the role assessment plays in language pedagogy (Fulcher, 2021; Levy-Vered & Alhija, 2015; Xu & Brown, 2016). This clearly suggests that teacher LAL, besides technical skills and principles required for assessment practices, should concern "a combination of cognitive traits, affective and belief systems" (Xu & Brown, 2016, p. 155) about assessment. The conceptions of LOA, therefore, encompass language teachers' belief systems about assessment knowledge where epistemological suppositions about the dynamic nature of assessment and learning are required to be possessed by teachers. This would foreground teachers' conceptualizations of assessment, their affective and cognitive responses to assessment approaches and types (e.g., assessment culture vs. testing culture), the role of assessments in society as well as their impacts on society, equity and fairness in assessment, assessment policies, purposes behind assessment, and decisions made on the basis of evidence obtained through assessment practices, among other factors (Fulcher, 2017, 2021). This implies that cognitive and affective aspects of language assessment as well as epistemological beliefs

are framed by teachers' conceptions of assessment (Brown, 2008). Furthermore, teachers' conceptions of assessment can influence teachers' implementation of their conceptual knowledge of assessment in practice (Watmani, et al., 2020), conceptual changes, and tendency to learn more about assessment.

Since almost half of the instructional time is devoted to various assessment-related tasks (Stiggins & Conklin, 1992), LAL can function as a yardstick to inform the evaluation of teachers' professional effectiveness and quality (Goldhaber, et al., 2015). However, which LAL skills and components to evaluate and how best to evaluate them to meet the changing assessment needs of the 21st century teachers, in both local or global contexts (Fulcher, 2012; Tsagari & Vogt, 2017; Vogt & Tsagari, 2014) have formed an ongoing debate of the time. Following different LAL models and definitions, several LAL measures are reported in the available literature on teacher LAL evaluation. The last few years have witnessed attempts to study these models and their various dimensions across different groups and contexts (Kremmel & Harding, 2020), however, the question relating to the way they are conceptualized and the extent they are possessed by a major group of language teachers, i.e., English as a Foreign Language (EFL) teachers, in specific localized contexts such as the one of the present study, has not fully touched yet. One reason is the scarcity of updated frameworks for exploring and measuring the *theoretical and practical knowledge base, personal beliefs, or affective dimensions* of teacher LAL across localized contexts. According to Fulcher (2021), the existing LAL must, therefore, extend to more than just assessment skills and techniques used in assessment practice to include teachers' conceptions, understanding, and beliefs about assessment, for instance, views and philosophy of language assessment.

Measuring teachers' LAL is important to judge their professional effectiveness and requires an appropriate and valid measurement tool (Liu et al., 2016; Stiggins, 1991). To date, there exist several instruments for tapping or measuring teachers' LAL base. However, they are mostly based on narrow conceptions of this trait where assessment is influenced by psychometric view emphasizing only technical or practical skills of language assessment. More domains and components were, however, recently echoed by researchers (e.g., Giraldo, 2018; Kremmel & Harding, 2020; Malone, 2013; Taylor, 2013; Xu & Brown, 2016) in the field. Fulcher (2012), for instance, pointed to assessment core domains in detail including abilities to "to place knowledge, skills, processes, principles and concepts within historical, social, political and philosophical framework" (Fulcher, 2012, p. 125) and years later, he (Fulcher, 2021) addressed LOA knowledge where epistemological suppositions about the dynamic nature of assessment and learning are required to be possessed by teachers. However, the question of such an assessment conceptualization, and other assessment needs and LAL levels of particular groups of language teachers, such as EFL teachers, still remains hypothetical. Besides, the pressing need for having such a wide range of competencies, concepts, and skills for performing appropriately and effectively in the current digital assessment era (e.g., theoretical and attitudinal/affective knowledge required for effective virtual assessment practices) makes it essential to develop LAL scales or update the existing ones to fulfill the peculiarities of today's local EFL assessment contexts in Iran.

Less empirically investigated is, therefore, what EFL teachers conceive and conceptualize as language assessment knowledge base, its nature, and purposes as well as its role in their pedagogy. To address this gap, the present study took into account the way EFL teachers in Iran conceptualize their LAL and what beliefs, attitudes, and values they hold towards it. Specifically, this study draws on Xu and Brown (2016) and Fulcher's (2012, 2021) 'conceptions of assessment' components in the LAL framework, as two of the recently proposed multi-componential guiding models for examining teachers' conceptions of theoretical knowledge base and its implementation in practice and focuses on the development and empirical examination of an update LAL instrument to help explore and identify teachers' conceptions, beliefs and the affective system underlying their overall assessment knowledge base. This, in particular, would illuminate the link between *theoretical* knowledge base, *belief*, and *affective* dimensions with *practical knowledge* dimension of assessment literacy.

Despite the pivotal importance of LAL for teachers, an adequate measure to assess both affective and theoretical dimensions of teachers' LAL with diverse assessment needs, such as EFL teachers, has not been fully developed. Evidence from studies on assessment literacy of EFL teachers indicates that they need to contextually apply their knowledge of assessment theory and technical skills to the practices of their classroom (e. g., Fulcher, 2012; Inbar-Lourie, 2013), however, many of them do not enjoy a high level of affective and theoretical literacy. More empirical studies on language teachers' assessment theoretical and affective knowledge are still required, more particularly in EFL contexts such as Iran where almost all teachers are non-native speakers and their LAL needs vary from other English language assessment contexts.

In line with the current global reform in language assessment and adoption of LOA in many language learning settings, EFL teaching context in Iranian schools reveals recent initiatives in reforming assessment practices and taking insights from current language assessment theories and views into the assessment of EFL students' achievement and learning (Author, 2020). This requires local EFL teachers to expand their LAL repertoire beyond mere technical skills and procedures and update their conceptual and epistemological beliefs and affective domains to perform a variety of classroom-based assessments. Inspired by Xu and Brown (2016) and Fulcher's (2012) view that an expanded definition of LAL should include not only the practice and principles but also the peculiarities of the assessment contexts, this study addressed Iranian EFL teachers' LAL through development and validation of a scale which can be used as a diagnostic tool for assessing teachers' affective and conceptual/theoretical competencies of their LAL. More specifically, the present study aimed to explore the extent to which the two hypothetically different dimensions (i.e., affective as well as theoretical/conceptual) of LAL are empirically distinct.

2. Review of Literature

Teacher assessment literacy is used as a yardstick in the evaluation of teachers' professional effectiveness (Stiggins, 1991). In general, assessment literacy refers to “an individual’s understandings of the fundamental assessment concepts and procedures deemed likely to influence educational decisions” (Popham, 2011, p. 267). The concept encompasses a set of

behaviors and competencies including teachers' ability to effectively administer assessments that target learning, score and understand learners' scores in a test, interpret learners' performance in light of a specific type of assessment, and communicate the accurate results to the stakeholders (Boyles, 2005; Stiggins 1999; Stoyhoff & Chappelle, 2005). Over the past years, teacher assessment literacy and its various competencies have been the focus of attention in assessment settings (Fulcher, 2021). Originated from general assessment literacy, LAL has been described as "the knowledge, skills, and abilities required to design, develop, maintain or evaluate, large-scale standardized and/or classroom based tests, familiarity with test processes, and awareness of principles and concepts that guide and underpin practice, including ethics and codes of practice and the ability to evaluate the role and impact of testing on society, institutions, and individuals" (Fulcher, 2012, p. 113). Similarly, in Taylor's (2009) elaboration, an assessment literate teacher is the one with an adequate understanding of language assessment principles, the ability to translate these principles into practice, select or develop appropriate assessment tasks, and make decisions about their students' status and performance outcomes.

Using survey data of 1086 stakeholders from different parts of the world to assess their LAL needs, Kremmel and Harding (2020) empirically examined Taylor's (2013) multidimensional framework. Their findings provided evidence in support of Taylor's LAL model including *technical skill, language pedagogy, impacts, and social values, knowledge of theories, principles and concepts, scoring, and personal beliefs/attitudes*. Similarly, LAL has been defined by Vogt and Tsagari (2014) as "the ability to design, develop and critically evaluate tests and other assessment procedures, as well as the ability to monitor, evaluate, grade and score assessments on the basis of theoretical knowledge" (p. 377). Underlying all of these LAL frameworks is a set of core components, values, and perspectives.

Designing an appropriate scale to assess teachers' LAL has been a hot topic in research since the 2000s (e.g., Bøhn, & Tsagari, 2021; Giraldo, 2018; Kremmel & Harding, 2020; O'Loughlin, 2013). Theoretical and empirical attempts have been made to assess teachers' LAL systematically through various measures, such as objective tests of assessment knowledge, surveys, or a set of rubrics. In fact, the earlier models were identified with traditional language testing and psychometric facets and contents. However, since the 2000s, as a consequence of increasing recognition of LOA culture, teachers' feedback literacy, assessment-informed teaching, and consequential effects of language assessments (e.g., Davies, 2008; Fulcher, 2012; Kane, 2012, 2013; McNamara & Roever, 2006; William, 2018), a redefinition of assessment literacy was proposed. For instance, Inbar-Lourie's (2013) systematic study on macro and micro views of LAL redefined it as a process-based framework including several components (pp. 27-41):

- Conceptions of current language assessment theories and methods, as well as views on language learning,
- Knowledge of technical skills, scoring, and interpretations,
- Proficiency in use of language assessment data (for decision-making and future instruction purposes),
- Mastery of exploring qualities of assessment tools (i.e., validity, reliability, practicality, the authenticity of assessment tasks), and

-
- Understanding of social dimension of language assessment.

Similarly, in Brown's (2008) view, LAL is argued to be framed and mediated by teachers' conceptions of assessment, including *affective* aspects (i.e., positive and negative emotions about assessment), *cognitive* aspects (i.e., what is true/false about assessment, or the in/consistency of teachers' new knowledge, skills and strategies with their current assessment conceptions and knowledge base), and *beliefs* about teaching and learning aspects. Both affective and cognitive dimensions as well as beliefs system are framed by personal, social, and educational experiences (Hill et al., 2010) and as mentioned before make teachers' conceptual change difficult or easy, in practice.

To date, one of the most comprehensive models of teacher literacy has been proposed by Xu and Brown (2016) who scaled teacher LAL along several core dimensions:

- The knowledge base (i.e., *disciplinary knowledge, knowledge of grading, knowledge of feedback, knowledge of assessment purposes, content and methods, knowledge of peer and self-assessment, knowledge of assessment interpretations and communication, and knowledge of assessment ethics*),
- The conceptions of assessment (i.e., *affective dimensions, views of learning and epistemological beliefs, and cognitive dimensions*), and
- The assessment literacy in practice (i.e., *assessment decision-making and action-taking*).

While teachers' conception of assessment has been acknowledged to play an undeniable role in shaping their knowledge about student assessment (Levy-Verd & Alhija, 2015), few studies have focused on it as a legitimate dimension of LAL (Xu & Brown, 2016). An awareness of these dimensions or competencies is crucially important for language teachers to develop their LAL and apply it appropriately in LOA settings.

Despite the multiplicity of labels used by contemporary LAL models, they show an improvement to the traditional conceptualizations of assessment literacy and tend to add affective and epistemological or personal aspects to teacher LAL to provide insights into the way teachers' LAL can be touched upon and measured in the current LOA era. However, most of the resulting measures are driven by earlier conceptions of language assessment standards such as measurement-only perspectives. Given the current conceptions of LOA requiring both teacher and student involvement in assessment practices, for instance, peer-assessment and self-assessment as argued by JCSEE (2015) and Xu and Brown (2016), the existing measures do not fully include the components or subscales required for identifying EFL teachers' conceptions of current assessment concepts and theories, epistemological beliefs, and affective facets (Looney et al., 2017; Pastore & Andrade, 2019). To address this gap, the present study aimed to develop and factor-analyze an updated LAL tool intended to explore how LAL conceptual/theoretical knowledge base and affective dimensions are realized in a local EFL assessment context of Iran where EFL teachers need to use a variety of assessments to be more accountable, accurately and appropriately, for student attainment and learning. Despite the paramount importance of LAL, empirical evidence for an updated scale measuring what these teachers conceive as assessment in the EFL domain is scarce.

Therefore, informed by the available literature, in particular, Xu and Brown (2016) and Fulcher's (2012) extended definition of LAL that yielded insights into LAL needs of specific

teachers practicing in specific contexts, the present study aimed at developing an appropriate instrument to adequately measure LAL levels of EFL teachers working in EFL contexts of Iran. It is argued that language teaching in such EFL contexts involves contextual peculiarities that can shape teachers' epistemology and beliefs about assessment, i.e., teachers' conceptions of assessment, (Latif, 2021). In addition to its development, dimensionality and appropriacy of the scale were targeted for a closer examination against the data obtained from the present EFL context. The following research question was thus investigated in this study:

To what extent are the hypothetical constructs assumed in a model of Iranian EFL teachers' LAL empirically distinct?

3. Method

3.1. Sample

To obtain a more comprehensive picture, the target population of this study comprised high school EFL teachers teaching grades 10-12 in public sector schools of several main geographical regions in Iran, including provinces in central regions (Tehran, Esfahan, and Yazd), south (Booshehr) and west (Azerbaijan). These are among the most highly populated provinces, have EFL teachers of different qualifications, and offer pre- and in-service teacher training programs. Both random and convenient sampling procedures were used to collect survey data from 200 EFL teachers, via in-person (110 paper surveys) and on-line forms (more than 100 electronic forms) due to the covid-19 pandemic condition.

The sample finally consisted of 213 EFL teachers (return rate=85%) working in educational sectors (i.e., public and private schools) of the selected provinces. 145 females (68%) and 68 males (32%) teachers, with the age range of 22-55 consented to take part in the study. Regarding their majors, 15 (7%) had their specialization in English literature, 8 (3.8%) in linguistics, 17 (8%) in translation studies and 173 (81.2%) in ELT (English language teaching). As to the level of education 70 (32.9%) had B.A. degree, 106 (49.8%) hold M.A., and 37 (17.4%) teachers were PhD. holders. With respect to assessment training courses or programs, 47 (22.1%) teachers reported 'no training', 80 (37.6%) teachers had 'pre-service training', 51 (23.9%) had 'in-service training' and the remaining 36 (16.4%) teachers reported both 'pre- and in-service training'.

3.2. Instruments and Procedures: Language Assessment Literacy (LAL) Scale

To collect teachers' conceptualizations of LAL, a self-report instrument was constructed. A comprehensive review of the literature was carried out for checking (a) available LAL models and frameworks, and (b) comprehensibility and intelligibility of the instrument items. This intensive review yielded more than 100 behaviors and themes (e.g., teachers' views on assessment consequences on learning; or aligning assessment objectives with national directives and policies) that were also detailed by Xu and Brown's (2016) hypothesized model. That is the *content analysis approach* taken for thematic analysis resulted into recurring themes that were verified by Xu and Brown's conceptions of the assessment framework, including disciplinary knowledge, assessment concepts and views of learning, showing conceptual dimension, and personal, social, and local values and beliefs, showing affective dimension. To brainstorm and develop initial items for these

scales, the relevant assessment literacy measures and models were consulted as well (see for instance, Bøhn, & Tsagari, 2021; Taylor, 2013; Vogt & Tsagari, 2014).

The initial pool of items was revised three times to check for the comprehensibility, relevance, and feasibility of the items, resulting into the following principal components that finally comprised the instrument: *Conceptual dimension*, including 1) disciplinary Knowledge (4 items), 2) assessment concepts, principles and purposes (6 items), and 3) assessment in language pedagogy (14 items); *Affective dimension*, including 1) personal beliefs and attitudes (3 items), and 2) social and local values (5 items).

Content and face validity of this 32-item Likert-type instrument were checked by five language assessment experts and two teachers. They were asked to state either agreement or disagreement with each item. Around 80% of these experts rated the items as relevant to teacher assessment literacy, revealing acceptable content analysis (see Appendix 1 for the final version of the survey instrument). Respondents could choose among four options: *not skilled at all*, *slightly skilled*, *skilled*, and *highly skilled*. Examples are how skilled/competent you are in 'providing students with a particular type of feedback corresponding to their particular error types?' (assessment in language pedagogy, item 27) or in 'knowing how my attitudes and beliefs might influence my assessment practices? (personal beliefs, item 46)'. The resulting version of the LAL scale was administered among 251 English language teachers; 213 were returned back (a response rate of 85%).

4. Results

To corroborate the factor structure of the designed LAL instrument, a series of factor analyses, both Exploratory (EFA) and Confirmatory Factor Analysis (CFA) were performed. First, to find empirical evidence for the dimensionality of the model and to explore these constituting factors, EFA analyses were conducted prior to CFA. To check the factorability of the data, Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) was run. The results (KMO= which is $>.05$) and a significant Bartlett's Test of Sphericity ($p <.05$) showed appropriacy of the factor analysis method.

All items were subjected to Principal Component Analysis (PCA) in EFA with a direct oblimin rotation (an oblique rotation). PAF was conducted to detect latent constructs underlying the model and oblimin rotation was used to detect the relatively high correlation between the emerging factors (see Pallant, 2010). The results showed that 7 factors with Eigenvalues exceeding 1 could be initially extracted, explaining 62.98% of the total variance. In addition, inspection of the scree plot also suggested that five factors should be retained because a very small change was observed between the values of 5 and 6 factors. Further analyses with different extraction modes were also re-run suggesting an empirical structure for the existence of five subscales (see Table 1). Items with loadings of $<.30$ were removed which led to the collection of 23 items. When the removed items were inspected closely degrees of similarity to another item were found (see Appendix 2 for the removed items). Of 32 items, 9 revealed very weak loadings on their factors.

Table 1.

Eigenvalues for five-factor solution

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	12.210	38.157	38.157	12.210	38.157	38.157
2	1.865	5.829	43.986	1.865	5.829	43.986
3	1.412	4.413	48.399	1.412	4.413	48.399
4	1.329	4.153	52.552	1.329	4.153	52.552
5	1.182	3.393	56.244	1.182	3.693	56.244

Following this step, the items making up the five main factors were examined for commonalities and developing or confirming labels. The first factor included 8 items each related to language pedagogy which was called *assessment in language pedagogy* (following Taylor, 2013; and Kremmel & Harding, 2020) including items indicating the use of assessment in learning and teaching contexts. The second factor contained 3 items, each related to *disciplinary knowledge (DK)* indicating teachers' knowledge of learning theories and trends, proficiency frameworks, and learner proficiency levels. *Assessment concepts, contents, and purposes* was the third factor which included five items relating to assessment terminologies, functions, principles such as fairness and ethics, and reliability, validity, and practicality concepts. The next two factors included items related to affective dimension, including factor 3 (*personal beliefs and attitudes*, as labeled by Kremmel & Harding, 2020) and factor 4 containing *local and social practices and values (social and local values)*.

While these five distinct factors emerged in exploratory mode, confirmatory factor analysis (CFA) with AMOS (version 22) was carried out to examine and confirm the factor structure (structural measurement) of this model, also determine the adequacy of the factor loadings and the standardized residuals that could not be obtained through EFA. All items were, therefore, subjected to separate CFA, using MLM (Maximum Likelihood Method). All items showed standardized factor loadings above .30 and thus retained in the scales. As Table 2 shows, all items loaded significantly on their designated factors and had acceptable factor loadings.

Examining the normed Chi-Square statistic (i.e., CMIN/DF) and other Goodness-of-Fit Indices (GFIs) showed a significantly fit model with CMIN/DF=1.39 (acceptable range of >1 & <3), Comparative Fit Index (CFI)=.96 ($\geq .90$), Goodness-of-Fit Index (GFI)=.91 ($\geq .90$), and RMEAS=.043 (<.08). Therefore, the 'Language Assessment Literacy Scale' showed a rather good-fitting model. The data were also checked for outliers, data points that exerted excessive leverage, and normality (Mardia's statistics). Statistics for skewness and kurtosis were also checked, and the results were below the suggested level of ± 2.0 , ranging from -.576 to .177, and -.578 to .573 for the variables, respectively, indicating no evidence of skewness or kurtosis.

Table 2 shows the elaborated estimation of the measurement model parameters achieved through CFA-AMOS (including loading estimates obtained from both EFA and CFA modes, standard errors, squared multiple correlations, and critical ratios).

Table 2.

Parameter Estimates of the Standardized Factor Loadings (CFA estimates), Standard Error (SE), Critical Ratio (CR), P-Value, and EFA estimates for the Measurement Model (LAL)

Items	5 Factors	S.E.	C.R.	P	Loadings	Loadings
					EFA estimates	CFA estimates
Q3	<--- Assessment_in Lang_Pedagogy			***	.445	.651
Q6	<--- Assessment_in Lang_Pedagogy	.108	7.650	***	.475	.593
Q7	<--- Assessment_in Lang_Pedagogy	.116	7.930	***	.659	.626
Q16	<--- Assessment_in Lang_Pedagogy	.124	7.924	***	.615	.621
Q17	<--- Assessment_in Lang_Pedagogy	.125	8.590	***	.761	.688
Q18	<--- Assessment_in Lang_Pedagogy	.126	9.262	***	.764	.745
Q19	<--- Assessment_in Lang_Pedagogy	.138	8.149	***	.783	.734
Q20	<--- Assessment_in Lang_Pedagogy	.124	7.756	***	.546	.606
Q10	<--- Disciplinary_Knowledge				.810	.627
Q30	<--- Disciplinary_Knowledge	.214	6.994	***	.533	.761
Q31	<--- Disciplinary_Knowledge	.103	10.981	***	.934	.679
Q21	<--- Assessment_C C & Purposes				.688	.623
Q22	<--- Assessment_C C & Purposes	.103	9.437	***	.719	.675
Q23	<--- Assessment_C C & Purposes	.125	7.623	***	.633	.659
Q24	<--- Assessment_C C & Purposes	.123	8.557	***	.728	.738
Q25	<--- Assessment_C C & Purposes	.120	6.499	***	.502	.515
Q27	<--- Personal Beliefs_& Attitudes				.717	.936
Q28	<--- Personal Beliefs_& Attitudes	.122	6.960	***	.476	.790
Q32	<--- Personal Beliefs_& Attitudes	.076	10.861	***	.935	.802
Q1	<--- Social_Local Values				.416	.453
Q5	<--- Social_Local Values	.192	5.843	***	.435	.542
Q9	<--- Social_Local Values	.201	5.492	***	.598	.485
Q26	<--- Social_Local Values	.225	6.318	***	.270	.655

Note: Assessment CC & Purposes=Assessment Concepts, Contents and purposes; Social Local Values=Social and Local Values ***=p-value<.01.

The analysis of items loading on the emerged factors, factor loadings, and p- values of the modified model could support the 5-component model of LAL conceptual and attitudinal dimensions. All items met the adequacy criteria at the CFA stage ($CR > 1.96$, $p\text{-value} < .05$, error variance, or $SE \leq 1.0$), therefore all 23 items of t EFA stage were kept in the model due to their significant loading on the LAL scale ($p < .05$). Table 3 presents the factors that were represented in the final version of the survey.

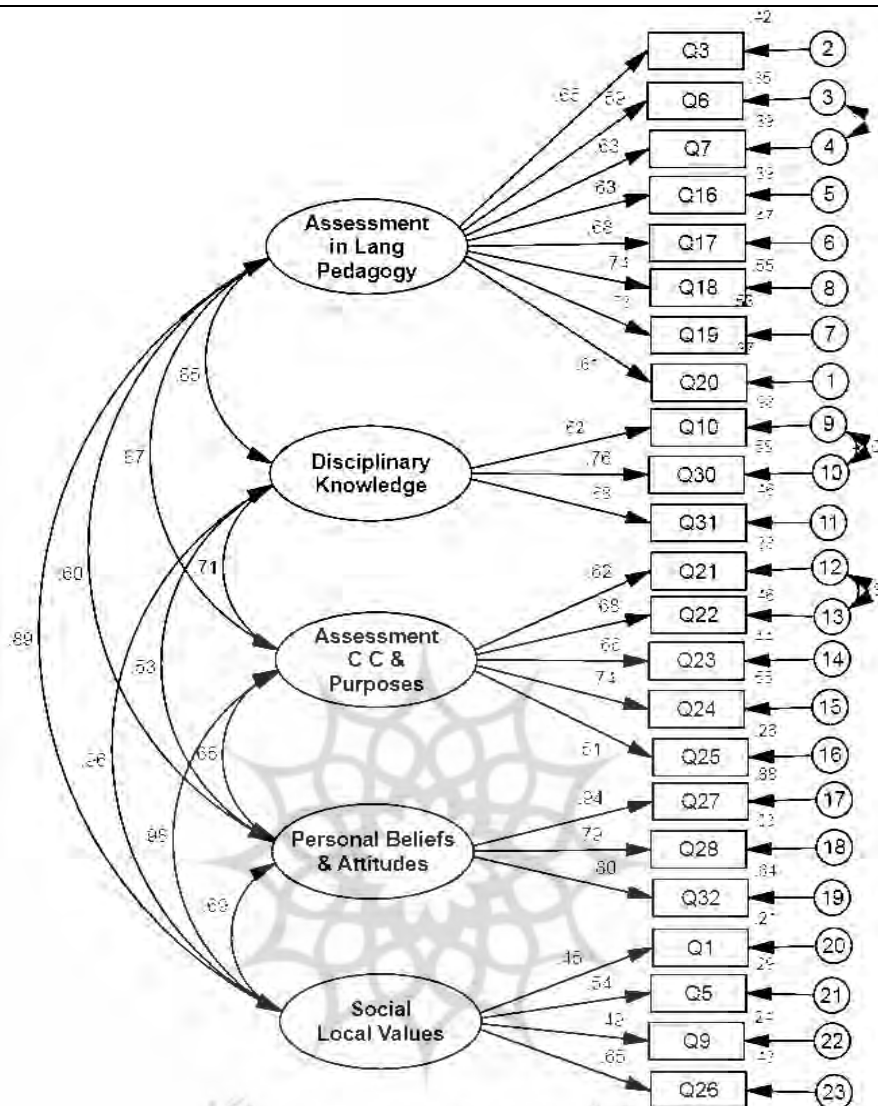


Figure 1. Final five-factor model of LAL (LAL Theoretical and Attitudinal dimensions)

As shown in Table 2, all loadings are above .30. The highest loadings were reported to be for factor 2 (*disciplinary knowledge*, e.g., I am competent in knowing different learning theories of linguistic competence (e.g., Universal grammar, Functional grammar), item 10), and factor 4 (*Personal Beliefs and Attitudes*, e.g. I am competent in knowing how my attitudes and beliefs might influence my assessment practices, item 46). The results also showed that the lowest CFA loadings belong to item 1 (*social and local values*: I am skilled/competent in determining if a classroom language assessment aligns with the policies of a national educational system).

Table 3.

Reliability coefficients of the LAL dimensions, item numbers, and Means

Teacher LAL scale and its factors	Items	\bar{x}	SD	α coefficient
LAL total	N=23	2.75	.464	.93
Assessment in language pedagogy	N=8 (3, 6, 7, 16, 17, 18, 19, 20)	2.81	.531	.86
Assessment theories	N=5 (21, 22, 23, 24, 25)	2.71	.542	.79
Disciplinary knowledge	N=3 (10, 30, 31)	2.83	.619	.78
Personal beliefs and attitudes	N=3 (27, 28, 32)	2.80	.620	.77
Social and local values	N=4 (1, 5, 9, 26)	2.65	.487	.70

The internal consistency of the total scale was found to be .93 for the whole scale and .86, .79, and .78 for assessment in language pedagogy (\bar{x} =2.81), assessment theories (\bar{x} =2.71), disciplinary knowledge (\bar{x} =2.83), respectively, showing satisfactory reliability indices. Rather acceptable reliability indices were also found for personal beliefs and attitudes (α =.77; \bar{x} =2.80) and social and local values (α =.70; \bar{x} =2.65) subscales.

5. Results

The present study proposes a contextualized LAL framework for examining Iranian EFL teachers' literacy of assessment in terms of their conceptions of language assessment. A synthesis of available literature on LAL models and thematic analysis proposed an initial LAL grid resulting into a self-assessment scale. The instrument invited EFL teachers to self-assess their conceptions of language assessment competencies and the LAL Levels they possessed. The factor structure of the scale was confirmed against the present data set (both paper-based and online) confirming the utility of the instrument for gauging EFL teachers' assessment knowledge base and affective competencies. The scale revealed our EFL teachers' conceptions of assessment are built on their understanding of assessment theories, concepts, and principles as well as beliefs and attitudes towards the contexts of their practice. This was empirically supported through a careful factor analysis of the scale, reducing the items down into five factors: (1) disciplinary knowledge, (2) assessment theories, (3) personal beliefs and attitudes, (4) social and local values, and (5) assessment in language pedagogy. The results, though with a reduction in the number of the extracted components, corroborate components of conceptions of assessment hypothesized by Xu and Brown's (2016) model, in particular, with 'disciplinary knowledge', 'knowledge of assessment purposes and theories', 'views of learning and epistemological beliefs' and 'attitudes and emotions'. The present findings are, furthermore, consistent with Taylor's (2013) 'knowledge of theories', 'principles and concepts', 'language pedagogy', and 'personal beliefs and attitudes'.

Specifically, three clearly emerging categories 'assessment in language pedagogy', 'assessment theories', and 'disciplinary knowledge' appear to relate to Xu and Brown's '*cognitive*' dimension, all referring to teachers' conceptual or theoretical knowledge base (taken as a higher-level factor). Secondly, 'personal beliefs and attitudes' and 'social and local values'

appear to be conceptualized as '*affective*' dimension as a higher-order factor of Xu and Brown's model. The practical implementation dimension showing what teachers perform in practice (i.e., assessment skills and scorings techniques) was not the focus of the present study, however.

In light of the present results, it can be argued that in Iranian EFL contexts, teachers have developed the literacy (in particular, disciplinary and pedagogical competencies) to deliver assessments that promote students' learning. The highly loaded items of disciplinary knowledge, and assessment principles and theories show that teachers have developed an understanding of LOA and its significance for their contexts. Even though the EFL assessment contexts, whether micro- or macro-, often place constraints on teachers' performance and autonomy required for quality assessment and informed decisions, their assessment expertise and knowledge base can be facilitated through professional development, for instance, through updating their LAL level, gaining new insights and holding clear conceptions of assessment. This representation puts conceptions of assessment at the heart of LAL. It also highlights the evolving and dynamic nature of teachers' LAL, a knowledge base which is not merely obtained through an accumulation of assessment experience but rather through the development of a set of contextually appropriate assessment conceptions and competencies (Xu & Brown, 2016). This professional knowledge, in turn, will contribute to teacher quality performance in various assessment contexts with special contextual idiosyncrasies, which seems to be the case with our local EFL contexts.

The highest loadings pertained to personal beliefs and attitudes items. Personal beliefs and attitudes may refer to how teachers' opinions and preconceptions may "inform [their] conceptualizations, judgments and decisions in assessment" (Scarino, 2013, p. 109). Our EFL teachers perceived their personal beliefs, preconceptions, and attitudes towards the assessment contexts highly important especially when almost all of them rated themselves as highly skilled in 'knowing that their attitudes and beliefs influence their assessment practice' (item 27, loading > .90). A relevant point here is that attention should be paid to teachers' attitudes and beliefs and to gearing them for optimal learning if LOA is adopted and expected to work efficiently (Figureas, 2021). Important with the affective dimension of LAL is that affective factors (e.g., beliefs, attitudes, motivation, and emotions) influence both assessment and learning success within LOA. There is empirical evidence indicating that learning outcomes are not merely affected by assessments but influenced by a psychological response such as perceptions and attitudes towards the assessment contexts (see for example, Bai, 2020; Wall & Horak, 2008; Murray et al., 2012). The present findings can, therefore, highlight the importance of affect dynamics for our EFL teachers and the need to address it in the development of a contextualized LAL.

Overall, the LAL factors and items endorsed by the present study are in line with the ones identified by some available studies. For instance, specific items such as 'knowing different language proficiency frameworks, e.g., the Common European Framework of Reference [CEFR], 'the ability to introduce peer- and self-assessment to students', 'the skill of conducting formative assessment', as well as the 'personal', 'social', and 'local' constructs converge with the similar themes described by Bøhn and Tsagari (2021) and Kremmel and

Harding's (2020) studies. The findings support the conception that translating assessment policies and principles into the realities of classroom practices requires language teachers' awareness and recognition of their own conceptions of assessment, specifically their literacy levels. The LAL model developed and confirmed in this study help meet this requirement. The scale, with its distinctive factors/constructs, contributes both to the conceptualization and consideration of LAL as a situated and dynamic knowledge system our EFL teachers constantly need to improve and use as a resourceful base to inform their LOA practices. In this way, classroom teachers, as assessment practitioners, can reach advanced literacy at the procedural level to extend their knowledge beyond ordinary concepts to include social, epistemological, and affective dimensions (Pill & Harding, 2013).

Since language teachers' assessment crafts build on their existing conceptualizations of assessment, self-interrogating their conceptions of assessment, in particular their personal conception of LOA, offers opportunities to reflect over their assessment practices, make subsequent changes in the interest of student learning, communicate accurate and accountable results to the stakeholders, and advance their LAL in practice (Inbar-Lourie, 2013; Pill & Harding, 2013; Stoyhoff & Chapelle, 2005; Xu and Brown, 2016). In other words, teachers' awareness and recognition of their own conceptions of assessment (i.e., both emotionally and conceptually) can empower them with autonomy and ownership of assessment practices. Despite these opportunities and resources, many teachers are often involved in making assessment decisions without sufficient conceptual and contextualized understanding of the nature and purposes of assessment (DeLuca & Klinger, 2010; Lam, 2015; Popham, 2009). Regarding the case with EFL teachers, for instance, research shows that they need to contextually apply their LAL theoretical, technical, epistemological, and disciplinary knowledge to their classroom practices (e. g., Fulcher, 2012) so they need to acquire such knowledge base, whether via in-service or pre-service training and workshops. Taken this view, the present study focused on the development and validation of a contextualized self-scrutiny or self-assessment LAL survey focusing on local EFL teachers' conceptions of LAL and its dimensions. As a prerequisite, therefore, teachers need to be aware of their own LAL level and improve it as a resourceful base to inform their assessment practices.

6. Conclusion

Building on insights from LOA and the literature on teacher LAL, this study aimed to present what comprises Iranian EFL teachers' conceptions and understanding of language assessment, reflecting teachers' epistemological, affective, cognitive, and conceptual knowledge base. To do so, a contextualized grid was initially developed, checked for its components and items, and finally empirically factor analyzed through both EFA and CFA processes, reducing the whole items down into five distinct components EFL teachers need to possess. The findings are in line with Xu and Brown (2016) in that teachers' conceptions and emotions about assessment approaches and paradigms (e.g., traditional testing paradigm vs. LOA) reveal a better understanding of teachers as both professionals and individuals, and that there is a pressing need for these individuals to get professionally developed and improve their LAL levels.

Overall, the findings bring EFL teachers' assessment literacy to the foreground, and draw much attention to the importance of teachers' personal conceptualization and emotions for the formation of their crafts and expertise in language assessment. Taking effective actions for enhancing students' learning, empowering them and achieving the intended learning goals, therefore, heavily relies on teachers' reconceptualization of LAL dimensions or competencies. Given the LOA purposes, a language assessment literate EFL teacher can employ different kinds of assessment competencies, such as the ones confirmed in this study, and use various assessment evidence to achieve measurably better EFL learning. This implies that EFL teachers' inadequate levels of assessment conceptions and emotions challenge their ability to assess and enhance their students' learning effectively, in practice.

The findings have various implications for the field. First, the present instrument can be used as an operationalized model for exploring 'teacher LAL in terms of 'conceptions of assessment concepts, principles, and disciplinary knowledge' as well as 'epistemological views', and 'emotions, beliefs, and attitudes towards assessments or assessment contexts'. Secondly, each of these distinctive but relevant components of the instrument can be used as a point for research across different contexts. For instance, when the prevailing views of language assessment in Iran (i.e., assessment for learning culture in some EFL contexts vs. exam-oriented culture in many others, especially public schools, see Authors, 2020) are concerned, finding from LAL studies can help explore the extent to which the LAL individual components or the whole components need to be contextualized in this particular L2 context. Thirdly, the results can be used in teacher training programs to (a) enhance EFL teachers' LAL, and (b) explore how such an assessment training can change them into a more assessment literate or, as acknowledged by DeLuca et al. (2013), a more learning facilitator.

The study suffered from some limitations, however. First, since the data were gathered during the Covid-19 pandemic (at the time of schools and institutes closures) we could obtain the survey responses only from 213 EFL teachers. A more representative EFL sample is required for scale development and validation like the one discussed here. Second, there might be some LAL items or categories that were not included in the survey items. This raises the need for using the experience of more researchers and scholars in the field. In another study recently done by the present authors (forthcoming), they developed LAL practical skills of Iranian teachers. These components together with the ones highlighted in the present study can be used together in a more inclusive contextualized model for local EFL teachers of the country. This could be done by future research.

Acknowledgments

The authors would like to thank all the teachers who participated in the study.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- Bai, Y. (2020). The relationship of test takers' learning motivation, attitudes towards the actual test use and test performance of the College English Test in China. *Language Testing in Asia*, 10(10), 1-18.
- Barnes, N., Fives, H., & Dacey, C. M. (2015). Teachers' beliefs about assessment. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teacher beliefs* (pp. 284-300). New York: Routledge.
- Bøhn, H., & Tzagari, D. (2021). Teacher educators' conceptions of language assessment literacy in Norway. *Journal of Language Teaching and Research*, 12(2), 222-233.
- Boyles, P. (2005). Assessment literacy. In M. Rosenbusch (Ed.), *National Assessment Summit Papers* (pp. 1-15). IA: Iowa State University. <https://doi:10.1.1.883.1970>
- Brown, G. T. L. (2008). Assessment literacy training and teachers' conceptions of assessment. In C., Rubie-Davies, & C., Rawlinson (Eds.), *Challenging thinking about teaching and learning* (pp. 285-302). New York: Nova Science.
- Carless, D. (2015). *Excellence in university assessment: Learning from award-winning practice*. London: Routledge
- Davies, A. (2008). *Assessing academic English: Testing English proficiency 1950-1989: The IELTS solution*. Cambridge, England: Cambridge University Press and Cambridge ESOL.
- DeLuca, C., & Klinger, D. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in education: Principles, Policies and Practice*, 17(4), 419-438.
- DeLuca, C., Chavez, T., & Cao, C. (2013). Establishing a foundation for valid teacher judgement on student learning: The role of pre-service assessment education. *Assessment in Education: Principles, Policy & Practice*, 20(1), 107-126.
- Figureas, N. (2021). Success in Language Learning: What Role Can Language Assessment Play. In A., Gebril (Ed.), *Learning Oriented Language Assessment: Putting theory into practice* (pp. 69-84). New York: Routledge.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132.
- Fulcher, G. (2017). The practice of language assessment. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 463-475). New York: Routledge.
- Fulcher, G. (2021). Language Assessment Literacy in a Learning-Oriented Assessment Framework. In A. Gebril (Ed.), *Learning Oriented Language Assessment: Putting theory into practice* (pp. 34-48). New York: Routledge.
- Gebril, A. (2021). Nationwide Assessment Reform in Egypt: Challenges and Potentials of Learning-Oriented Assessment in an Examinations-Based Instructional Setting. In A. Gebril (Ed.), *Learning Oriented Language Assessment: Putting theory into practice* (pp. 109-122). New York: Routledge.

- Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile: Issues in Teachers' Professional Development*, 20(1), 179-195. <https://doi:10.15446/profile.v20n1.62089>
- Goldhaber, D., Lavery, L., & Theobald, R. (2015). Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students. *Educational researcher*, 44(5), 293-307. <https://doi:10.3102/0013189X15592622>
- Hamp-Lyons, L. (2017). Language assessment literacy for learning-oriented language assessment. *Papers in Language Testing and Assessment*, 6(1), 88-111.
- Hill, M. F., Cowie, B., Gilmore, A., & Smith, L. F. (2010). Preparing assessment-capable teachers: What should pre-service teachers know and be able to do? *Assessment Matters*, 2, 44-64.
- Inbar-Lourie, O. (2013). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385-402.
- Joint Committee on Standards for Educational Evaluation (JCSEE). (2015). *The need for student evaluation standards*. Retrieved from <http://www.jcsee.org/wpcontent/uploads/2009/09/SESNeed.pdf>
- Kane, M. (2012). Validating score interpretations and uses. *Language Testing*, 29(1), 3-17.
- Kane, M. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50, 1-73.
- 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.
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169-197. <https://doi:10.1177/0265532214554321>
- Latif, M. W. (2021). Exploring tertiary EFL practitioners' knowledge base components of assessment literacy: Implications for teacher professional development. *Language Testing in Asia*, 11(1), 1-22. <https://doi:10.1186/s40468-021-00130-9>
- Levy-Vered, A., & Alhija, F. N.-A. (2015). Modelling beginning teachers' assessment literacy: The contribution of training, self-efficacy, and conceptions of assessment. *Educational Research and Evaluation*, 21(6), 378-406.
- Liu, S., Xianxuan, X., & Stronge, J. S. (2016). Chinese Middle School Teachers' Preferences regarding Performance Evaluation Measures. *Educational Assessment, Evaluation and Accountability* 28(2), 161-177. <https://doi:10.1007/s11092-016-9237-x>
- Looney, A., Cumming, J., van Der Kleij, F., & Haris, K. (2017). Reconceptualizing the role of teachers as assessors: Teacher assessment identity. *Assessment in education: Principle, Policy & Practice*, 25(5), 442-467.
- McNamara, T. & Roever, C. (2006). *Language Testing: The social dimension*. Oxford: Blackwell Publishing.
- Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329-344. <https://doi:10.1177/0265532213480129>

- Murray, J., Riazi, M. & Cross, J. (2012). Test candidates' attitudes and their relationship to demographic and experiential variables: the case of overseas trained teachers in NSW, Australia. *Language Testing*, 29(4), 577-595.
- Pallant, J. (2010). *SPSS Survival Manual: A step by step guide to data analysis using SPSS*. McGraw Hill.
- Pastore, S., & Andrade, H. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, 84(2), 128-138.
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381-402.
<https://doi:10.1177/0265532213480337>
- Popham, W. J. (2009). Assessment Literacy for Teachers: Faddish or Fundamental? *Theory Into Practice*, 48, 4-11. <https://doi:10.1080/00405840802577536>
- Popham, W. J. (2011). *Everything school leaders need to know about assessment*. Thousand Oaks: Corwin Press.
- 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.
<https://doi:10.1177/0265532213480128>
- ShayesteFar, P. (2020). A Model of Interplay between Student English Achievement and the Joint Affective Factors in a High-stakes Test Change Context: Model Construction and Validity. *Educational Assessment, Evaluation and Accountability*, 23(3), 335-371. <https://doi:10.1007/s11092-020-09326-8>
- ShayesteFar, P., & Fazlali, F. (2020). Raising the Stakes of High School Exit Exams: Students' Perspectives of their English Learning Motivation in a High-Stakes Test Change Context. *International Journal of Language Testing*, 10(2), 71-87.
https://www.ijlt.ir/article_118015
- Mohammadkhah, E., Kiany, G. H., Tajeddin, Z., & ShayesteFar, P. (2022). Language Teachers' Conceptions of Assessment Literacy Theories and Skills: A Contextualized Measure for the Ministry of Education Teachers in Iran. *Journal of Teaching Research Quarterly*, X, XX-XX.
- Stiggins, R. (1991). Assessment literacy. *The Phi Delta Kappan*, 72, 534-539.
- Stiggins, R. J. (1999). Assessment, student confidence, and school success. *The Phi Delta Kappan*, 81(3), 191-198.
- Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of classroom assessment*. Albany, New York: SUNY Press.
- Stoynoff, S., & Chapelle, C. A. (2005). *ESOL tests and testing*. Alexandria, VA: TESOL.
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21-39. <https://doi: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:10.1177/0265532213480338>

- 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. <https://doi.org/10.1080/15434303.2014.960046>
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374-402. <https://doi.org/10.1080/15434303.2014.960046>
- Wall, D. & Horak, T. (2008). *The impact of changes in the TOFEL examination on teaching and learning in central and eastern Europe: phase 2, coping with change*. ETS, TOEFL.
- Watmani, R., Asadollahfam, H., & Behin, B. (2020). Demystifying Language Assessment Literacy among High School Teachers of English as a Foreign Language in Iran: Implications for Teacher Education Reforms. *International Journal of Language Testing*, 10(2), 129-144. https://www.ijlt.ir/article_118025
- William, D. (2018). *Embedded formative assessment* (2nd ed.). Bloomington: Solution Tree Press.
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149-162. <https://doi.org/10.1016/j.tate.2016.05.010>

Appendix 1: Teacher Language Assessment Literacy (LAL) Scale

(find the complete version in the link:

https://docs.google.com/forms/d/e/1FAIpQLSfEo5oWy0RTKK_nYrJSX6CBI0ZWBiBOYEMR_qoGRCceAH4g/viewform?usp=sf_link

Dear colleague,

Please specify the extent you perceive you are skilled/competent in the assessment domains described by the following statements.

(please choose one of the Not skilled at all=1; slightly skilled=2; skilled=3; highly skilled=4)

Domains <i>How skilled/competent you are in.....?</i>	LAL Level			
	Not skilled at all 1	Slightly Skilled 2	Skilled 3	Highly skilled 4
1 determining if a classroom language assessment aligns with the policies of a national educational system.				
2 knowing how major language teaching approaches and methods (e.g. CLT, TBLT) affect selection of language assessment methods (e.g. dynamic assessment).				
3 knowing how major language learning theories and trends (e.g. Cognitivism, Socio-cultural theories, CLA, ...) affect selection of assessments trends (e.g., Integrative, and Functional-pragmatic trends).				

-
- 4 understanding if the local context of teaching might influence language assessment in my classes (e.g. when group work is/not the norm in a local area).

 - 5 aligning assessment objectives (e.g., instructional/learning objectives) with national directives and policy documents (e.g. Document of Educational Reform) during specification and designing phases.

 - 6 knowing how assessments can influence the design of a language course or curriculum.

 - 7 knowing the consequences of assessments on classroom teaching, learning and materials.

 - 8 recognizing students' levels of language proficiency (e.g. A1 to C2 reading, listening, writing, speaking levels) according to standard frameworks (e.g. CEFR).

 - 9 knowing the influence of social values (e.g., individualism) on language assessment practices (i.e., designing, implementing or using language assessment tasks).

 - 10 knowing different learning theories of linguistic competence (e.g., Universal grammar, Functional grammar...).

 - 11 knowing different forms of alternative assessments (e.g. collaborative tests, take-home tests, portfolios, ...).

 - 12 how to provide useful feedback to improve students' learning.

 - 13 providing various types of feedback (e.g., metalinguistic, direct, indirect) based on students' performance on assessments.

 - 14 knowing different student responses to assessment-related feedback (e.g. revision required, attention to correction only).

 - 15 providing students with a particular type of feedback corresponding to their particular error types.

 - 16 applying technology or digital assessment tools (e.g. mobile apps, web-based- platforms) to give feedback to students.

 - 17 training my students how to self-assess their performance.

 - 18 training my students how to peer-assess each other's performance.

 - 19 engaging my students in self-monitoring their learning by using assessment information.

 - 20 knowing how to statistically evaluate language assessments regarding their reliability or validity.

 - 21 identifying if an item on a test offends or unfairly penalizes students simply because of race, gender, religion or socioeconomic status (i.e., assessment bias).

 - 22 maintaining equal assessment protocols for all the students involved (e.g., equal time, allocation, ...).

 - 23 providing individualized learning opportunities to meets students' different needs.

 - 24 knowing if assessment outcomes are used inappropriately (e.g., unfair pass/fails and selection decisions).
-

-
- 25 accommodating candidates with disabilities or other learning impairments.
-
- 26 recognizing if the assessment content is appropriate (e.g., culturally, socially, locally).
-
- 27 Knowing how my attitudes and beliefs might influence my assessment practices.
-
- 28 understanding if I have attitudinal conflicts with other assessment stakeholders (parents, administrative staff).
-
- 29 understanding the impacts that language assessments/tests have on students' lives and promotion (e.g., getting a job in future).
-
- 30 knowing different language proficiency frameworks (e.g., the Common European Framework of Reference [CEFR]).
-
- 31 knowing different learning theories accounts of whats and hows of language learning or acquisition.
-
- 32 perceiving my personal attitudes towards language assessment.
-

Appendix 2: list of removed items

No.	Item
2	knowing how major language teaching approaches and methods (e.g. CLT, TBLT) affect selection of language assessment methods (e.g. dynamic assessment).
4	understanding if the local context of teaching might influence language assessment in my classes (e.g. when group work is/not the norm in a local area).
8	recognizing students' levels of language proficiency (e.g. A1 to C2 reading, listening, writing, speaking levels) according to standard frameworks (e.g. CEFR).
11	knowing different forms of alternative assessments (e.g. collaborative tests, take-home tests, portfolios, ...).
12	how to provide useful feedback to improve students' learning.
13	providing various types of feedback (e.g., metalinguistic, direct, indirect) based on students' performance on assessments.
14	knowing different student responses to assessment-related feedback (e.g. revision required, attention to correction only).
15	providing students with a particular type of feedback corresponding to their particular error types.
29	understanding the impacts that language assessments/tests have on students' lives and promotion (e.g., getting a job in future).