

## The Status Quo of Language Assessment Literacy among Syrian EFL Teachers

Duaa Sulaiman<sup>1</sup>, S. Susan Marandi<sup>2\*</sup>, Leila Tajik<sup>3</sup>

<sup>1</sup> PhD Candidate, Department of English Language and Literature, Literature Faculty, Alzahra University, Tehran, Iran

<sup>2</sup> Associate Professor, Department of English Language and Literature, Literature Faculty, Alzahra University, Tehran, Iran

<sup>3</sup> Assistant Professor, Department of English Language and Literature, Literature Faculty, Alzahra University, Tehran, Iran

Received: 2020/11/22

Accepted: 2021/05/23

**Abstract:** With the increasing importance of Language Assessment Literacy in recent years, identifying the assessment literacy components in different contexts becomes essential to ensure that language teacher professional development is on the right path, thus providing students with proper evaluation. This study describes a research project in which an adapted version of Fulcher's (2012) *Language Assessment Literacy Survey* was delivered via the Internet in an attempt to characterize the levels of assessment knowledge of English language teachers in Syria. Three hundred and thirty Syrian English language teachers participated in the study. Exploratory and confirmatory factor analyses were applied to the data obtained from the constructed-response item, and qualitative data analysis procedures were applied to the open-response items. The results indicated that Language Assessment Literacy in the Syrian context mainly comprises four factors: the social impact of tests; test prepping and administration; test design, development, and interpretation; and evaluating language tests. On the other hand, the content analysis applied to the responses given by Syrian English teachers to the open-response questions indicated failings in both theoretical and practical assessment literacy in Syria's education scene, particularly the former, calling for an immediate change in teachers' preparatory courses in Syria. It is hoped that the study results can help language teacher education programs specify EFL teachers' academic assessment goals and enhance the nature of future language assessment programs, particularly in Syria. It can further provide a basis for comparison with other contexts and countries, thus contributing to a cross-cultural understanding of language assessment literacy.

**Keywords:** Language Assessment Literacy, Language Teacher Education, Factor Analysis.

\* Corresponding Author.

Authors' Email Address:

<sup>1</sup> Doa'a Soleman (doasoleman90@gmail.com), <sup>2</sup> S. Susan Marandi (susanmarandi@alzahra.ac.ir), <sup>3</sup> Leila Tajik (tajik\_l@alzahra.ac.ir)

ISSN (Online): 2322-5343, ISSN (Print): 2252-0198

© 2021 University of Isfahan. All rights reserved

## Introduction

### *Language Assessment Literacy (LAL)*

The term Language Assessment Literacy (LAL) appeared following the emergence of the term Assessment Literacy (AL) in general education (Stiggins, 1991). According to Edwards (2017) “Assessment literacy is an ongoing process that requires continuous teacher improvement through initial teacher education (ITE), pre, and in-service teacher preparatory courses, and ongoing teaching experience” (p. 2). Being assessment literate means that teachers need to obtain the knowledge required to assess students’ language proficiency and to know how to interpret the results and use them to improve their teaching instructions (Boubris & Haddam, 2020).

LAL can be defined as the familiarity of language instructors with assessment tips and the techniques needed to evaluate the language proficiency of students (Kim, Chapman, Kondo, & Wilmes, 2020). Despite the notable impact of assessment on teaching and learning processes alike, some studies argue that many EFL teachers are not well literate in different assessment themes (Koh, Burke, Luke, Gong, & Tan, 2018). Many pre-service and in-service teacher training programs do not contain testing/assessment practical preparation (Watmani, Asadollahfam, & Bahram, 2020). As Stiggins (2014) rightly said, “we cannot continue to turn a blind eye to practitioners' lack of competence in classroom assessment” (p. 72). However, LAL has not been defined precisely so far because of its situated status, which is unique for each context; and this unclear identity of LAL requires further investigations in future studies (Coombe, Vafadar, & Mohebbi, 2020).

Mere theoretical knowledge is not enough; instead, a comprehensive and contextual application of the theoretical concepts in educational contexts can improve teachers' assessment levels (Koh *et al.*, 2018). It has become vital to provide qualified training programs for teachers in order to help them in designing their adapted testing samples to support students' learning (Coombe *et al.*, 2020). Little assessment knowledge caused by inadequate professional academic assessment training makes English teachers reluctant to make testing decisions independently as well as incapable of composing valid and dependable classroom exams. To overcome this mismatch, improving Language Assessment Literacy is crucial to promote students’ learning (Boubris & Haddam, 2020).

### *Language Assessment Literacy in Context*

The status of language in society and learners' education causes variations in the components that feature LAL in different contexts. Specifically, LAL cannot be alienated from its surrounding context because it is a constructive interpretive approach to the surrounding professional background (Coombe *et al.*, 2020). However, this diversity leads to a lack of the training necessary for teachers to apply the standard assessment practices for their contexts and to enhance the teaching pedagogy and learners' achievement (Watmani *et al.*, 2020). In addition, the growing impact of the Common European Framework of Reference (Council of Europe, 2001) in the current century has advocated the position of assessment in line with teaching and learning processes. As a result of these modern advancements in the educational curriculum, content, and instruction, in-service as well as pre-service teachers need to continuously follow the recent professional assessment improvements (Vogt, Tzagari, & Csépes, 2020). However, the humble literature published so far attempting to practically define each context's assessment components further extends the difficulties and complexities (Yan & Fan, 2020). Therefore, there is consensus that more research on Language Assessment Literacy is required to introduce a transparent, situated assessment literacy framework within language testing to define the needed type of assessment literacy for any professional program (Jan-nesar, Khodabakhshzadeh, & Motallebzadeh, 2020).

Language assessment, as a separate discipline, is not systematically recognized in Syria and very little research exists in this regard (Syrian Ministry of Education and scientific research, n.d.). In a rare exception, Mohamad, Sarma, and Mohapatra (2018) explored the washback effect of the Syrian National English test of the Baccalaureate on the classrooms' teaching instructions; however, no research has been done to detect the components of English teachers' assessment literacy in the Syrian context until now. This is while in order to help EFL teachers improve their language assessment literacy, we have to collect information on their current language testing knowledge (Tavassoli & Farhady, 2018). The objective of this study, therefore, is to investigate the LAL of English language teachers in Syria; the results can be used by policymakers and teacher educators to organize and develop more effective EFL teacher education programs in the future in Syria. Thus, the present study aimed to answer the following research question:

What is the status quo of the Language Assessment Literacy (LAL) of Syrian EFL teachers as represented via an adapted version of Fulcher's survey (2012)?

## Review of the Related Literature

### *The role of Language Assessment Literacy (LAL) in Improving Teachers' Educational Competencies*

LAL, an indispensable part of Assessment Literacy (AL), enhances the language classroom quality and learners' achievement (Edwards, 2017). Assessment literate EFL teachers are capable of situating the selected assessment practices to match the needs of learners (Scarino, 2013). As such, EFL teachers need to update their theoretical and practical assessment knowledge/skills continuously to improve their teaching and students' learning. Teachers ought to know how to prepare detailed scoring criteria and acknowledge the mechanisms needed to record the scoring results for more correct grade-based decisions that influence test-takers' academic and social life (Tzagari & Vogt, 2017).

To improve education, teachers need adequate training programs in assessment (Islam, Hasan, Sultana, Karim, & Rahman, 2021). As a result of modern developments in the educational curriculum, content, and instruction, in-service and pre-service teachers need to continuously follow the new professional advancements of assessment (Watmani *et al.*, 2020). Despite the important role of testing in the educational context, many educational colleges do not urge pre-service teachers to enroll in a specified classroom assessment course, which negatively influences their teaching practices (Janatifar & Marandi, 2018). Qualified training programs to assist practitioners in designing their adapted testing practices should be implemented in different education contexts to promote students' learning (Koh *et al.*, 2018).

In language assessment, experts call to administer assessment programs that educate EFL teachers on the standards needed to select and improve suitable assessment techniques and to score, interpret, and validate the results to make reasonable educational decisions (Nimehchisalema & Bhattib, 2019). As Popham (2004) put it, ignoring assessment literacy is a "professional suicide" on part of the teachers. Indeed, LAL should be explored further in the literature addressing different educational contexts (Coombe *et al.*, 2020). This calls for an investigation of the status quo of LAL of teachers in the Syrian context, which is the focus of the present study.

### *English Language Assessment in Syria*

The English language has received increasing interest in Syria within the last two decades due to its status as an international language of science and communication. After the 2002

education reform in Syria, teaching English was promoted to the main subject starting from first grade, and taught by non-native language teachers who were expected to use the textbooks '*English for Starters*' in seven hours of weekly classes (Hos & Cinarbas, 2017). Along with the promoted position of the English language, testing English in the Syrian educational system increased widely in the last six years, leading to three-fold school and college English exams compared with the average number of tests held prior (Mohamad *et al.*, 2018). While this test expansion has helped improve the teaching quality and learning environment, it has caused some novel challenges for stakeholders, test developers, and EFL teachers, especially since language assessment is usually not part of the training of Syrian language teachers, nor does substantial research exist in this regard.

Nevertheless, the researchers of this study perused the electronic library of the doctorate and master's degree theses of the Syrian Ministry of Higher Education and scientific research (MOHE) via the URL: <http://mohe.gov.sy/mohe/index.php?node=5714>, to ensure the novelty of this topic in Syria. They found neither doctorate nor master's degree university studies related to language assessment in the Syrian context (Syrian Ministry of Higher Education and scientific research MOHE, n.d.), except for one study by Mohamad *et al.* (2018), which has explored the relationship between Syrian EFL teachers' assessment literacy and the washback effect in preparing for the standardized exams. Moreover, a proposal entitled 'The Influence of Formative Assessment on EFL Speaking Proficiency' was registered among the MA thesis titles in TEFL documented in the Higher Institute of Languages in Damascus University (2019). Therefore, it seems safe to say that to date there is no research study done on the components of the AL in the Syrian context.

On further investigating the context with regard to language assessment education, the following related points were noted: The higher institutes of languages in both Damascus and Latakia both grant a Master's degree in Teaching English as a Foreign Language which requires a thesis (Higher Institute of Languages in Damascus University, 2021; Higher Institute of Languages in Tishreen University, n.d.). On the other hand, similar institutions in Homs and Aleppo grant an MA degree in TEFL, yet require no thesis (Appendix 3) (Higher Institute of Languages in Al-Baath University, n.d.; Higher Institute of Languages in Aleppo University, n.d.). However, the Higher Institute of Languages also offers a Diploma Programme in ELT that includes an evaluation course in its second semester.

According to the portal of the Higher Institute of Languages, the language assessment courses in the two granted certificates (i.e. MA, and Diploma) address some theoretical

themes of language testing without paying attention to the practical issues, or recent practices (Higher Institute of Languages in Damascus University, 2021; Higher Institute of Languages in Tishreen University, n.d; Higher Institute of Languages in Al-Baath University, n.d). For this reason, Syrian EFL teachers have inadequate exposure to the concepts and practices of classroom assessment in their national assessment training programs, which are not more than a few sessions about practical examples of some experienced teachers in their pre-service training programs. Thus, instead of following functional contextual assessment themes, Syrian EFL teachers resort to their teaching experience to write the items/questions of their exams (Mohamad et al., 2018). EFL teachers' status as instructors and assessors at the same time is complicated due to the fact that the main requirement to teach English in Syrian schools is holding a certificate in English Literature. According to Damascus University (2011), the English literature certificate provides no assessment course throughout the four years of courses. Therefore, an overwhelming majority of EFL teachers in Syria are not well-trained to adapt assessment procedures to their context.

Contrary to the language assessment literacy level of school EFL teachers, exams are given an increasing interest and are being extensively administrated to Syrian students, who are required to pass two English exams each semester. Additionally, two national standardized English exams are held at the end of the mandatory pre-secondary, and upper secondary schools; the second is the all-important Baccalaureate Examination which determines their university field (Mohamad *et al.*, 2018). In its endeavor to improve the quality of assessment in Syria, the Ministry of Education (MOE) has adopted the Common European Framework of Reference standards (Hallak *et al.*, n.d.). Following this, they recently issued a question template, a kind of test specification formula, for grades seven to twelve. Syrian teacher trainers and education advisors who monitor teachers through regular visits to schools insist on teachers following the designed templates to compose adequate exams (Hallak *et al.*, n.d.). However, the social pressure related to the standardized exams has made many Syrian EFL teachers fall into the trap of teaching to the test so that their students score good final marks instead of focusing on actually improving their language proficiency.

## Methodology

To answer the research question, a mixed-methods research design was used, combining quantitative and qualitative data.

### *Participants*

To answer the research question in this study, and to run exploratory factor analysis, 330 EFL school teachers who studied at different universities in Syria completed an adapted version of Fulcher's (2012) LAL survey (Appendix 1). Convenience sampling which is about reaching the most easily accessible members of the target population to the researcher (Etikan, Musa, & Alkassim, 2016) was used in the study. Meyers, Gamst, and Guarino (2013, p. 687) suggest, "To run exploratory factor with a 25-item inventory, you better have at least 300 participants." The participants took part willingly in the study, being fully aware of the fact that their data will be used for research purposes; nonetheless, pseudonyms are used to protect their anonymity. The demographic information of the participants appears in Table 1.

**Table 1.** *Demographics of the Participants*

	Categories	N
Gender	Male	90
	Female	240
Age	21-30	34
	31-40	267
	41-50	26
	51-60	2
	61-70	0
	Missing	1
Education Level	High School graduate	4
	BA	274
	MA	37
	PhD	1
	Other	14
	Missing	0
University	Damascus University	121
	University of Aleppo	54
	Tishreen University	94
	Al- Baath University	48
	Al-Furat University	11
	University of Hama	2
	Missing	0
Teaching Experience	0-5 years	67
	5-10	227
	10-20	26
	More than 20 years	10
	Missing	0

### *Data Collection and Analysis*

The researchers used an adapted electronic version of Fulcher's (2012) Language Assessment Literacy (LAL) Survey to explore the level of familiarity of EFL teachers studying and working in Syria with twenty-three assessment and testing topics. In order to answer the research question, some items were deleted from the original, such as "*Which is your home country?*" as all the participants were Syrian EFL teachers, as well as questions such as "*When you last studied language assessment, which parts of your course you thought [sic] were most relevant to your needs?*" since language assessment courses are in general not offered in Syria. Other items were altered; for example, instead of asking "*Which was the last language testing book you studied or used in class?*" the participants were asked *whether* they had ever read or used a language testing book. In addition, one of the original questions was designed to learn how important certain language assessment topics were held to be by the teachers, and the items listed were ranked accordingly (on a Likert-scale from *unimportant* to *essential*). Upon piloting the test, it was recognized that many Syrian teachers were unfamiliar with some of the concepts and thus unable to respond. We, therefore, changed the question to indicate how familiar they were with each topic; thus, the labels of the closed-response item choices were also changed to range from *not at all proficient* to *highly proficient*. Finally, the formatting of some questions was revised to suit the survey's electronic design for practicality purposes. Previous studies in the literature (e.g., Fulcher, 2012; Janatifar & Marandi, 2018; Tavassoli & Farhady, 2018) indicated the suitability of Fulcher's (2012) Language Assessment Literacy Survey for the purpose of representing the language assessment literacy of EFL teachers in different contexts. In fact, the Cronbach value of this survey was 0.93 in Fulcher (2012) and 0.83 in Janatifar and Marandi (2018), which are highly acceptable values. In addition, Cronbach's alpha obtained in the current context was a relatively high value of 0.75. The validity of the survey for the addressed context was further verified through an Exploratory Factor Analysis (EFA), achieved via the Statistical Package for Social Sciences (SPSS, version 22) and Confirmatory Factor Analysis (CFA) via the Analysis of Moment Structure (AMOS, version 24).

A major part of the results of the study pertains to question two of the survey, the closed response item that explores the familiarity of Syrian EFL teachers with some testing topics. This question was measured on a 5-point scale from *not at all proficient* to *highly proficient*, as pointed out above, and was statistically analyzed. Questions one and four are of a qualitative nature and inquire about the perspectives of Syrian EFL teachers concerning the



testing skills they need and essential topics for testing books. The last part of the survey addresses the respondents' demographic information and contains eight closed-response items (i.e., questions three, five, six, seven, eight, nine, ten, & eleven). The survey results were collected online from August 2019 to April 2020 from 330 ELT teachers in Syria, in the hope of exploring the assessment knowledge of Syrian EFL teachers and reaching a language assessment base for the Syrian educational experience. Before running the analysis with IBM SPSS, the researchers checked all the factor analysis assumptions, including normality, linear relations, factorability, and sample size (Meyers, Gamst, & Guarino, 2016).

### ***Context***

This study aimed to investigate the LAL of English teachers who teach and work at Syrian public schools. Language testing as a subject is not taught at Syrian universities (Damascus University, 2011), thus it is only to be expected that as far as education is concerned, Syrian English teachers are lacking in critical literacy concerning assessment. So, the present investigation of the language assessment knowledge among EFL teachers at public schools in Syria is largely indicative of their classroom experience.

### ***Design of the Study***

The researchers used a mixed-methods approach using the adapted LAL survey which contains items of a quantitative and qualitative nature. The multidisciplinary approach of language evaluation entails applying multiple research methods to reach some reasonable amount of information for different stakeholders (Riazi & Candlin, 2014).

### **Results and Discussion**

The research question was answered in part through factor analyses using the data gathered from the closed-response item (item 2) in the modified survey. As mentioned earlier, in this question respondents defined their level of familiarity with twenty-three testing terms according to their classrooms on a 5-point scale from *not at all proficient* to *highly proficient*. All the factor analysis assumptions, including normality, linear relations, homoscedasticity, independence of errors, and sample size, were checked before performing the analysis. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were applied to the data. The exploratory phase determined the underlying factor structure of LAL; the

confirmatory phase emphasized the suitability of the observed relationship between the factors found in the exploratory analysis.

The responses were entered into the Statistical Package for Social Sciences software (SPSS, version 22) to apply exploratory factor analysis. The analysis identified four LAL factors with their means and reliability estimates. In the confirmatory Factor Analysis phase, the results were entered into the statistical software Analysis of Moment Structures (AMOS 24) to analyze the adequacy of the model drawn between factors extracted at the EFA. The reliability of the scores from the survey's second question was found to be .75, which is an acceptable value (Meyers et al., 2016).

The researchers further explored the research question via the responses of the constructed-response items in the adapted survey, that is, items one, three, and four (Appendix 1). Question number one was about the skills needed for the Syrian context; question number three asked Syrian English teachers about their testing reading background. Finally, the researchers asked the participants about the content of a good testing book based on their language testing experience. These three questions were analyzed qualitatively in light of the EFL teachers' familiarity with the language testing themes, as reported by the four factors found in the exploratory phase. These factors explained the recurring issues of language assessment in the Syrian classrooms in a coding matrix design, which altogether constituted the final LAL model evaluated in the AMOS software.

### ***Defining the Model***

The research question was analyzed using the participants' responses to questions 1, 2, 3, and 4. Exploratory Factor Analysis (EFA) examined the participants' answers to question 2 of the adapted survey (Appendix 1). As mentioned earlier, all factor analysis assumptions were initially checked. EFA is about determining the variables that identify the latent factors based on a theoretical rationale (Meyers et al., 2016). Factor analysis helps test development and test scoring research to validate or even organize the format of a test or a survey or the measures used in a research program (Quaigrain & Arhin, 2017). The Cronbach's alpha in the current study was .75 (compared to .93 in Fulcher's research and .83 in Janatifar & Marandi's study). A value of .70 or above is suitable to estimate the appropriateness degree of the factor analysis' correlations (Kaiser, 1970; Kaiser, 1974). The Kaiser-Meyer-Olkin measure of sampling adequacy was also suitably high (.75). As mentioned earlier of both Cronbach's alpha and Kaiser-Meyer-Olkin measure of sampling adequacy, the values were within an

acceptable range. For data analysis in EFA, the researchers applied the Promax rotation (Maskey, Feib, & Nguyen, 2018).

According to the EFA results, language assessment literacy in the Syrian context can be broadly perceived as comprising the following four extracted factors: *the social impact of tests*; *test prepping and administration*; *test design, development, and interpretation*; and *evaluating language tests*. The first factor labeled *the social impact of tests* included the following assessment issues: deciding what to test, the uses of tests in society, side effects of the test on teaching, large-scale testing/national tests (9<sup>th</sup>/12<sup>th</sup> grades), and classroom assessment. The second factor, (i.e. *test prepping and administration*) comprised: preparing learners to take tests, educational measurement principles, test administration, and use of statistics. The third factor, *test design, development, and interpretation*, included: writing test specifications/ blueprints, procedures in language test design, rating performance tests (speaking, writing), interpreting scores, and selecting tests for use. Finally, the last factor, *evaluating language tests*, comprised reliability, evaluating language tests, and history of language testing.

The EFA results are found in Appendix 2, with the items listed in the left-hand column. As the table shows, the eigenvalues which emerged for the four factors were all greater than one, accounting for 43.446 % of all the constant variance: 15.545% for factor one, 25.801 % accumulatively for factors 1-2, 34.868 % for factors 1-3, and 43.446 % for all four factors. Based on the EFA results, the reliability and descriptive statistics for the four factors obtained in the current study are as follow in Table 2.

**Table 2.** Reliability and Descriptive Values for the Four Factors Based on the EFA Results

Factor	<i>Cronbach's α</i>	M	SD	SE
The social impact of tests	.74	3.182	.70820	.03899
Test prepping and administration	.80	2.995	.84109	.04630
Test design, development, and interpretation	.72	3.225	.77283	.04267
Evaluating language tests	.82	2.920	.93623	.05185
Total	.75			

It is worth noting that the results of the present study were somewhat different from those of Fulcher (2012) and Janatifar and Marandi (2018). In Fulcher's study (2012), the results suggest that testing cannot stand without three pillars: knowledge, skills, and principles in terms of both classroom and standardized assessment. Similarly, in Janatifar and Marandi (2018), Iranian EFL teachers appeared to believe that testing knowledge depends on teachers' theoretical and practical knowledge to appropriately evaluate their students. Their perception of the necessity of hands-on skills-based instruction in language assessment was in line with their theoretical background knowledge of testing. However, in the current study, the humble knowledge of Syrian EFL teachers of the testing terms presented to them made them incapable of providing a solid theoretical rationale for their assessment conceptions and information. Their knowledge depends almost solely on their experience concerning the assessment tasks required for their classrooms. As mentioned earlier, language testing is not a subject taught at university education, nor even in pre-service and in-service preparatory courses, which prepare EFL teachers mainly for teaching, except for some very humble assessment tips added recently for teachers' education institutes (Higher Education Ministry in Syria, n.d). Therefore, this makes teachers who studied at Syrian universities lack both the theoretical background of testing as well as an organized, practical application of the theoretical knowledge. In fact, due to the novelty of what language assessment knowledge represents for English teachers in Syria, the researchers were obliged to modify the second question of Fulcher's survey (2012) to detect the language testing knowledge in Syria by exploring the familiarity level of English teachers with testing terms. Unlike Iranian EFL teachers who reported weakness in practice compared to their theoretical knowledge, English teachers in Syria have inadequate theoretical knowledge of assessment influencing the students' evaluation and depend wholly on their experiential knowledge and more experienced teachers instead of a supportive theoretical rationale. In the current study, the findings and emerging themes suggest that Syrian EFL teachers need adequate language testing education to help them appropriately improve and implement different assessment practices in their classrooms. EFL teachers in Syria need to receive both practical and theoretical instructions concerning language assessment, particularly the latter.

This is further evidenced through a comparison between the results of the factor loadings of the present study with those of Fulcher (2012) and Janatifar and Marandi (2018). Fulcher (2012) had sent the survey to EFL teachers who resided in different countries yet had benefited from language testing education. Similarly, Janatifar and Marandi (2018) sent the

survey to Iranian EFL teachers who had already learned about language assessment and were capable of providing useful feedback about their assessment knowledge and needs. The assessment information gleaned from the English teachers who had participated in these two studies showed only minor differences in the item loadings of the extracted factors. However, the present study dealt with staff untutored concerning language assessment, making the item loadings clearly different from the two above-mentioned studies. These differences can be observed in Table 3.

**Table 3.** Comparison between Janatifar and Marandi (2018), Fulcher (2012), and the Present Study

	<i>Janatifar &amp; Marandi (2018)</i>	<i>Fulcher (2012)</i>	<i>The present study</i>
<b>Factor 1 Items</b>	Test design and development D, F, E, H, I, G	Test design and development D, F, E, C, B, M	The social impact of tests C, V, P, O, S
<b>Factor 2 Items</b>	Large-scale standardized testing and classroom assessment Q, P, N, O, R	Large-scale standardized testing Q, P, N, H, G, V, W, L	Test prepping and administration R, T, W, L
<b>Factor 3 Items</b>	Beyond-the-test aspects V, U, W, C	Classroom testing and washback S, R, O, U, T, I	Test design, development, and interpretation D, B, G, M, I
<b>Factor 4 Items</b>	Validity and reliability J, K, L	Validity and reliability J, K	Evaluating language tests J, F, A

### *Evaluating the Model*

After checking various possible factor structures, the present four-factor structure model (i.e., the social impact of tests; test prepping and administration; test design, development, and interpretation; and evaluating language tests) emerging from the EFA was deemed to have the most interpretable results and was checked for the goodness of fit before proceeding to do a CFA. Next, in order to better evaluate the model obtained in EFA, the researchers applied

Confirmatory Factor analysis (CFA). CFA resembles a tool whose role is to reject or emphasize a measurement model (Tomé-Fernández, Fernández-Leyva, & Olmedo-Moreno, 2020). A prominent feature concerning CFA is its hypothesis-dependent characteristic concerning the model structure that contains specific factors underneath some items. When performing the analysis, the covariance value between the items is estimated to evaluate the hypothesized factor structure. Based on an *a priori* hypothesis, the researchers tested the concluded model's compatibility in reflecting the shape of the current concluded data set statistically (Alavi *et al.*, 2020).

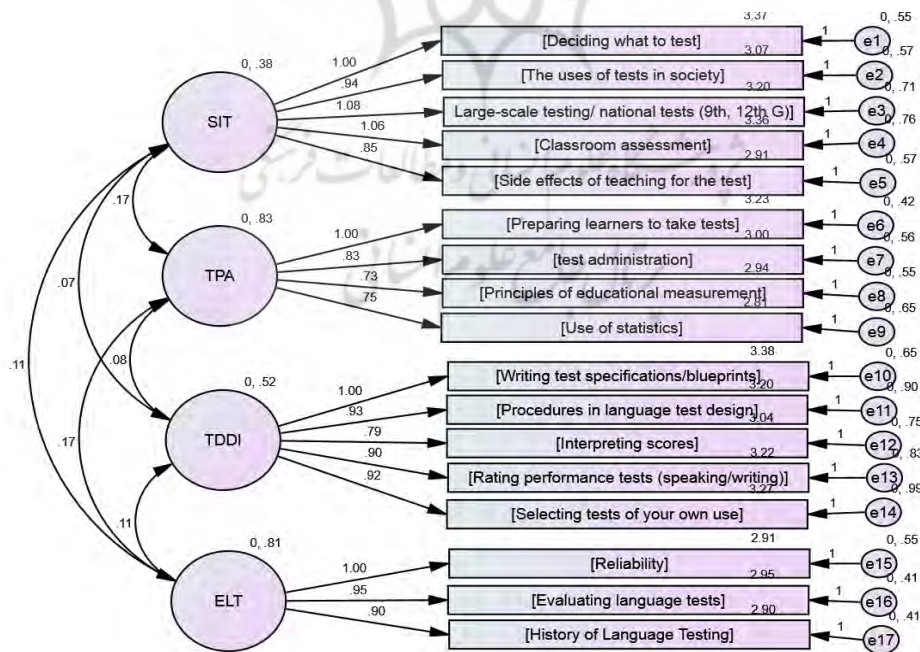
The assumptions for CFA were inspected meticulously. CFA can be conducted with positive degrees of freedom (Meyers *et al.*, 2016), among other requirements and assumptions which vary in different contexts (Kline, 2011). Since a one hundred percent fit is not possible in real-life settings, researchers aim to detect the model's relative level of fit. One indicator of fit is a non-significant Chi-square value (Kline, 2005). In addition, the value obtained from dividing the Chi-square results by the degrees of freedom should preferably be less than two (Alavi *et al.*, 2020).

There are quite a variety of statistical tests in CFA to ensure a model fit, yet there is no general agreement on which is the preferred one; thus, no one method is universally adopted (Klem, 2000). Besides the Chi-square, other indices such as the p-value, which ought to be insignificant, the Root-Mean-Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), the Normed Fit Index (NFI) in line with others like the Akaike information criterion (AIC), the Browne-Cudeck criterion (BCC), and the expected cross-validation index (ECVI) should be mentioned (Meyers *et al.*, 2016). Nevertheless, some scholars have also suggested reporting the Tucker Lewis index (TLI) or Non-normed fit index (NNFI) with a value higher than 0.90 (Moss, 2014); others do not advise it since it is similar to the NFI (Meyers *et al.*, 2013). Each fit index reports a value as an indicator of the suitability of the model concluded. Both the NFI and the CFI should report a value of .95 for a good fit model. RMSEA should be less than .06 to indicate a good fit (Tomé-Fernández, *et al.*, 2020). The results obtained for some of the goodness-of-fit indicators in the study's CFA phase can be found in Table 4.

**Table 4.** Fit Indices of the Four-Factor Model of Language Assessment Literacy (LAL)

Model	$\chi^2$	df	$\chi^2/df$	p	RMSEA	CFI	TLI	NFI	AIC	BCC	ECVI
	94.433	133	.836	.897	0.000	1.000	1.019	.936	208.433	215.031	.634

For this study, the Chi-square calculated at this stage was significant [ $\chi^2 (113) = 0.906$ ,  $p > .05$ ]. One reason for this  $p$ -value is that  $\chi^2$  value is sensitive to the sample size (Alavi *et al.*, 2020). Besides this result of  $\chi^2$ , the calculated value for  $\chi^2/df$  was  $.836 < 5$ , which indicates a good fit (Meyers *et al.*, 2016). Moreover, the estimated coefficients of the indicator variables are statistically significant, suggesting that they are indicators of their respective factors (Kenny, 2020). Without a stable theoretical justification, no modifications can be proposed to the existing model, which makes the above-mentioned four-factor model the final model of language assessment literacy in the Syrian context, despite the usual slight difficulties in naming some of the factors due to the seeming incongruence of some items loading on certain factors, such as the loading of items L (i.e., use of statistics) and W (i.e., principles of educational measurement) on factor 2, or the loading of item A (i.e., history of language testing) on factor 4. In addition, items E, H, K, N, Q, and U did not load on any of the four factors. The resulting four-factor structure model of LAL among Syrian EFL teachers may be seen in Figure 1.



**Figure 4.** Confirmatory Factor Analysis (CFA) Results

Despite having a low assessment literacy background and due to their experiential knowledge, Syrian EFL teachers tended to believe that they have adequate familiarity level with the identified testing themes and appeared to feel that their assessment knowledge is relatively good (76.6%). However, they were more engaged with activities related to the first factor (the social impact of tests) in class, namely, deciding what to test, the uses of tests in society, side effects of teaching for the test, large-scale testing/national tests, and classroom assessment. Overall, the teachers perceived their assessment knowledge of item C (deciding what to test) to be higher than their recognition and acknowledgment of the other elements. Out of the four items loaded on the second factor (test prepping and administration), more than half of the participants (56.7) showed a high familiarity level with item R (preparing learners to take tests), most likely due to the social pressure of the surrounding context on both teachers and learners to gain high scores. However, such practices could indicate poor teaching and inauthentic learning for the sake of merely passing the test and scoring high (Mohamad *et al.*, 2018). English teachers were reasonably familiar with item T (test administration), and generally believed they have a higher knowledge degree of item W (principles of educational measurement) than item L (use of statistics). Indeed, the participants had comprehended item L as the general overall calculations of the semester-exams and class activities, rather than the statistical assessment trends of evaluating a test use, exploring reliability and validity of test scores, and reporting the results using different statistical tests. This might help explain why this item loaded on the second factor (Test preparation and administration) instead of the third factor (Test design, development, and interpretation).

In general, the third factor (Test design, development, and interpretation), which comprised *writing test specifications/ blueprints, procedures in language test design, rating performance tests, interpreting scores, and selecting tests for your own use* had the lowest familiarity levels for English teachers who work at Syrian public schools. These items require an assessment literacy education base that is not available for Syrian English teachers since it is not part of their academic education. Accordingly, EFL teachers in Syria build their assessment acquaintance on their classroom experiences and examination protocols of their classrooms and education institution.

Regarding factor four (Evaluating language tests), the results showed surprisingly high loadings of all three items (i.e., reliability, evaluating language tests, and history of language testing), with the values: .847, .842, and .820, respectively (See Appendix B). The high



loading of item A (history of language testing) on this factor was particularly surprising and somewhat inexplicable.

A point worth mentioning was that Syrian English teachers' teaching experience influenced their responses more than their education level; specifically, the more the teaching years, the more the familiarity with situating different classroom assessment practices. For the first factor, teachers who had been teaching from five to ten or more years were more familiar with aspects related to deciding what to evaluate following the institution's requirements for grade levels (54.7%), the national standardized exams (48.4%), and classroom assessment and evaluation practices (62.3%). On the other hand, teachers who had been teaching less than five years evaluated their acquaintance level with the same testing aspects as relatively low and requiring further both education and instruction (30.2%). The novice teachers in Syria informed the researchers that they tend to ask and follow the experienced teachers' assessment tips.

Similarly, regarding the second factor, Test prepping and administration, teachers with high experience years ranging from five to ten or more years showed a modest familiarity with item D (writing test specifications/ blueprints). This is while EFL teachers in Syria do not prepare or design any test specification templates; instead, the Ministry of Education has recently issued test specification templates for all the school grades (seventh grade till twelfth grade) to be adopted and followed by English teachers in writing, timing, and scoring their English exams (Hallak et al., n.d). Nonetheless, the experienced EFL teachers reported adequate assessment familiarity with the different language test design procedures (40%), which can be imputed to their classroom experiences concerning classroom tasks.

On the other hand, even experienced teachers with five or more years of teaching and classroom work reported low proficiency levels in interpreting scores (33.9%) and selecting tests for their own use (29.2%). These results further indicate the importance of improving Syrian EFL teachers' assessment education (pre-service training), and extending this education to those currently teaching at Syrian public schools (in-service training).

As stated above, the education system in Syria doesn't provide assessment training in preparation programs and university education, although general assessment topics are now being introduced in the first year of the Master's degree in Teaching English as a Foreign Language at the Higher Institute of Languages in both Damascus and Latakia. The evaluation course is also offered in the second semester of the 'Diploma Programme in ELT' at the same institutions, but it does not appear to provide updated assessment material that can adequately

improve teachers' LAL to enhance the classroom's education quality. Accordingly, education levels did not distinguish between EFL teachers in Syria concerning the familiarity level since the provided assessment knowledge is quite humble and addresses general tips which require updating. For the first factor, more than half of the BA level participants (54.7%) with more than five years of teaching experience estimate a high familiarity degree concerning 'deciding what to test', while MA-level participants recorded less familiarity (39.2%). Likewise, the item 'the uses of tests in society' (item V) was more familiar to EFL teachers with higher experience, irrespective of their educational background. While BA level participants (40.7%) showed higher familiarity level and knowledge of the uses of tests in their own social contexts, MA students with fewer experience years reflected their low level of familiarity with the uses of tests for classroom purposes or a specific institutions' intended targets. Again, education background did not help the teachers better identify classroom testing themes. Teachers with five to ten or more years of teaching experience (59.85%) showed a high familiarity level concerning classroom assessment procedures (item O) and the preparation, writing, and understanding of the test formula of the national standardized English school exams in Syria which are the ninth, and twelfth grades (item P). On the other hand, MA teachers who had just started teaching (37.5%) had a low familiarity level with the selection and appropriateness of the best assessment practices for their classes (item O), and the content of the large-scale standardized exams in Syria (item P). Despite having learned about these topics, many EFL teachers were unfamiliar with them in practice and resorted to obtaining help from more experienced teachers.

This is not to imply that experience can substitute the role of assessment education; rather, it emphasizes the insufficient and unsatisfactory nature of the current education practices. No doubt improving the assessment literacy level of Syrian EFL teachers would lead to making their classroom assessment practices more adequate and helpful in meeting students' demands based on their context and needs. For instance, it was observed that while MA participants who had actually experienced some form of testing education in their studies acknowledged having a medium familiarity level with writing and using test specification templates (item D) for a test (43.4%). They showed a low familiarity level concerning evaluating speaking and writing skills individually (item M) unless there is a rating rubric to follow (30%). One of the primary purposes of having adequate assessment literacy lies in adequate score interpretation and decision making. However, with the poor assessment literacy and education background of EFL teachers in Syria, EFL teachers with MA degrees

and low experience years did not differ significantly from those with lower educational backgrounds or other study fields. Around thirty percent of the MA participants claimed familiarity with interpreting students' competence based on their performance on their classroom exams in a specified content, and a close 28.7 % of the English teachers with low education background and teaching experience were familiar with the interpretation of students' scores. Regarding the fourth factor, approximately half the teachers holding an MA expressed a familiarity with item A (i.e., history of language testing; 51.2%), and almost half were also familiar with item J (reliability; 45.7%); on the other hand, only 25% of them believed they were proficient in evaluating language tests (item F). Unfortunately, there were no participants with PhDs among the EFL teachers who responded to the survey, and while this is perhaps telling in itself, it hindered the researchers from investigating their opinions and familiarity level with the identified testing topics.

In order to complete the picture obtained through the factor analyses, the researchers further applied content analysis to the open-response items of the survey, in which the Syrian EFL teachers gave their opinions regarding their required testing skills and knowledge. This was done by codifying the data into themes based on the already specified testing topics identified in the second question. The open-ended items that were used for the content analysis were items one (i.e. assessment tips and skills required for the Syrian EFL classrooms) and four (i.e. their opinions about a good testing book's content). The answers given by EFL teachers to the open-ended questions of the survey were categorized and compared with the four factors structure obtained at the factor analysis phase of the study.

Based on this analysis, 55% of the respondents considered the items contained in the first factor (the social impact of tests) as highly required topics within their education context; accordingly, they should be accorded more prominence in Syria's academic preparation programs for teachers. The second factor, Test prepping and administration, was evident in 20% of the responses. The third factor, 'test design, development, and interpretation', was mentioned by 19% of the respondents, and the remaining 6% of English teachers who responded to the open-ended questions emphasized the necessity of learning how to evaluate language tests. Based on the findings, most participants insisted on the necessity of implementing language testing education as part of teachers' preparatory courses. They emphasized including both the practical aspects of language testing as well as their theoretical justifications. Such a request by Syrian EFL teachers results from the absence of language testing materials at the university level, with very few assessment classes at

teachers' preparatory institutes (Damascus University, 2011). According to Boubris and Haddam (2020), the success of the teaching process is linked principally to the assessment protocol adopted in the class, making the need to improve the LAL of teachers' preparatory courses a necessity for the success of any educational context. On a side note, the researchers also noted that certain language testing issues such as 'procedures in language test design' and 'classroom assessment' were highly emphasized by teachers of varying degrees of experience, educational background, as well as different ages, whereas certain other concepts, such as 'test administration' and 'writing test specifications/ blueprints', only received more attention by teachers with higher education, especially those few who had actually experienced a language testing course.

The humble assessment literacy of the Syrian EFL teachers scrutinized in the current study indicates an urgent need to add specialized language testing courses to the teachers' preparatory courses, whether at the university or institutes. English teachers' language assessment literacy in a troubled education system such as that of Syria, which suffers from the war crisis, is a necessary preliminary for addressing English learners' low language proficiency level, which has been exacerbated by the current situation. To seriously improve the assessment literacy of Syrian English teachers, language assessment education needs to be addressed in a systematic and considered manner to ensure a much-needed enhancement of the education process of the Syrian classrooms. This is in line with Inbar-lourie's (2017) insistence on the importance of contextually-related assessment practices in testing courses.

## Conclusion

This study aimed at presenting a framework of LAL in the Syrian context; in other words, shedding light on the status of language assessment literacy in Syria. The researchers used a modified version of Fulcher's (2012) survey with two types of closed and constructed response items to explore language testing's status quo among Syrian teachers. According to the EFA and CFA results, the Language Assessment Literacy of Syrian school teachers comprises four factors: the social impact of tests; test prepping and administration; test design, development, and interpretation; and evaluation of language tests. Syrian EFL teachers appeared to be more familiar with themes that relied on their practical experience, but were severely lacking in the theoretical knowledge of assessment, which could naturally influence their classroom assessment practices (Syrian Ministry of Education and scientific research, n.d.). Paying due attention to Syrian EFL teachers' LAL needs in both pre- and in-

service teacher education programs can only result in enhancing the education system, and it is hoped that the results of this study will serve as a springboard for revitalizing LAL in Syria. The present study is also hoped to provide a foundation for comparison to other contexts, in order to achieve a cross-cultural understanding of LAL.

The current study had some limitations commonly found in such kinds of research. Firstly, the respondents were volunteers, thus they were likely to be those already interested in the topic. Another issue is the ever-present possibility that the participants said what they thought they should say rather than what they actually believed. Another limitation is that the researchers were not actually in Syria at the time of the study despite the fact that the first researcher is Syrian; thus, the interactions with the Syrian participants took place via the Internet. The last limitation is related to both the numerical and qualitative data in the survey. The quantitative results are naturally influenced by the topics and content included in the survey. By the same token, the qualitative data, which allowed the participants more freedom to express themselves by reflecting on some conceptual categories and descriptive themes, can also be said to be influenced by similar views through the researchers' own tinted lenses.

## References

- Alavi, M., Visentin, D. C., Thapa D. K., Hunt G. E., Watson R., & Cleary, M. (2020). Chi-square for model fit in confirmatory factor analysis. *Wiley*, 76(9). 1-3.
- Boubris, A. A., & Haddam, F. (2020). Reading Assessment: A Case Study of Teachers' Beliefs and Classroom Evaluative Practices. *Arab World English Journal*, 11(4), 236-253.
- Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: What do we need to learn, unlearn, and relearn?. *Language Testing in Asia*, 10(1), 1-16.
- Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. United Kingdom, UK: Cambridge University Press. Retrieved from: <https://rm.coe.int/16802fc1bf>.
- Damascus University. (2011). *English literature college Education plan*. Retrieved October, 10, 2020, from [http://new.damascusuniversity.edu.sy/faculties/humanscience/pdf/english\\_plan.pdf](http://new.damascusuniversity.edu.sy/faculties/humanscience/pdf/english_plan.pdf).

- Edwards, F. (2017). A rubric to track the development of secondary pre-service and novice teachers' summative assessment literacy. *Assessment in Education: Principles, Policy & Practice*, 24(2), 205-227.
- Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of Convenience Sampling and Purposive Sampling, *American Journal of Theoretical and Applied Statistics*, 5(1), 1-4.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132.
- Hallak, H., Al-thiab, L., Bayazid R., Sadek, B., Dawood Agha, O., Nasser A.,... Moazen, M. (n.d.). *Assessment Guide Grades 7-12*. (Ed.). Syrian Arab Republic Ministry of Education: National Centre for Curriculum Development.
- Higher Education Ministry in Syria (n.d.) <http://www.mohe.gov.sy/mohe/>. The electronic library of the doctorate, and master degree theses. Retrieved May, 10, 2021, from <http://mohe.gov.sy/mohe/index.php?node=5514&cat=4688&>.
- Higher Institute of Languages in Al-Baath University (n.d.). Retrieved May, 1, 2021, from <http://institute.albaath-univ.edu.sy/li/>
- Higher Institute of Languages in Al-Baath University (n.d.). *TEFL MA material description*. Retrieved May, 1, 2021, from <http://institute.albaathuniv.edu.sy/li/article292&phpMyAdmin=52cc71c91116d60b4b159c7b1761bb0d#.Xbx1A1RKjIV>
- Higher Institute of Languages in Aleppo University (n.d.). Retrieved May, 1, 2021, from <https://www.facebook.com/HIOL.ALEPPO/>
- Higher Institute of Languages in Damascus University (2021). Retrieved May, 10, 2021, from <http://damascusuniversity.edu.sy/arabicD/>
- Higher Institute of Languages in Damascus University (n.d.). *TEFL MA material description*. Retrieved May, 10, 2021, from <http://damascusuniversity.edu.sy/arabicD/?lang=1&set=3&id=378>
- Higher Institute of Languages in Damascus University (n.d.). *Theses' titles, TEFL MA*. Retrieved May, 1, 2021, from <http://damascusuniversity.edu.sy/arabicD/?lang=1&set=3&id=371>
- Higher Institute of Languages in Tishreen University (n.d.). Retrieved May, 1, 2021, from <http://www.tishreen.edu.sy/ar/high-institute/humanities/linguistics>

- Higher Institute of Languages in Tishreen University (n.d.). *TEFL MA material description*. Retrieved May, 7, 2021, from <http://www.tishreen.edu.sy/ar/high-institute/humanities/linguistics>
- Hos, R., & Cinarbas, H. I. (2017). Education interrupted: English education policy from the Rubble in Syria. In *English Language Education Policy in the Middle East and North Africa, language Policy* (13, pp. 223-234). Springer, Cham. Retrieved from [https://link.springer.com/chapter/10.1007/978-3-319-46778-8\\_13](https://link.springer.com/chapter/10.1007/978-3-319-46778-8_13).
- Inbar-Lourie O. (2017). Language Assessment Literacy. In: E. Shohamy, I. Or, & S. May (Eds.) *Language Testing and Assessment. Encyclopedia of Language and Education* (3rd ed. pp.257-270). Springer International Publishing. [https:// DOI 10.1007/978-3-319-02261-1\\_19](https://doi.org/10.1007/978-3-319-02261-1_19).
- Islam, M. S., Hasan, M. K., Sultana, S., Karim, A. & Rahman, M., M. (2021). English language assessment in Bangladesh today: principles, practices, and problems. *Language Testing in Asia*, 11(1). 1-21.
- Janatifar, M., & Marandi, S. S. (2018). Iranian EFL Teachers' Language Assessment Literacy (LAL) under an Assessing Lens. *Applied Research on English Language*, 7(3), 361-382.
- Jan-nesar, M. Q., Khodabakhshzadeh, H., & Motallebzadeh, K. (2020). Assessment Literacy of Iranian EFL Teachers: A Review of Recent Studies. *Journal of Asia TEFL*, 17(2), 689-698.
- Kaiser, H. F. (1970). A second-generation Little Jiffy. *Psychometrika*, 35(4), 401–415.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36.
- Kenny, D. A. (2020). *Measuring model fit*. Retrieved July 4, 2020, from <http://davidakenny.net/cm/fit.htm>.
- Kim, A. A., Chapman, M., Kondo, A., & Wilmes, C. (2020). Examining the assessment literacy required for interpreting score reports: A focus on educators of K–12 English learners. *Language Testing*, 37(1), 54-75.
- Klem, L. (2000). Structural equation modeling. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and Understanding More Multivariate Statistics* (pp. 227–260). Washington, DC: American Psychological Association.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.
- Koh, K., Burke L., E. C. A., Luke A., Gong W., & Tan C. (2018). Developing the assessment literacy of teachers in Chinese language classrooms: A focus on assessment task design. *Language Teaching Research*, 22(3), 264-288.
- Maskey, R., Fei, J., & Nguyen, H. O. (2018). Use of exploratory factor analysis in maritime research. *The Asian journal of shipping and logistics*, 34(2), 91-111.
- Meyers, L. S., Gamst, G., & Guarino, A. J. (2013). *Applied multivariate research: Design and interpretation* (2<sup>nd</sup> Ed.). SAGE Publications.
- Meyers, L. S., Gamst, G., & Guarino, A. J. (2016). *Applied multivariate research: Design and interpretation* (3<sup>rd</sup> Ed.). SAGE Publications.
- Mohamad, M., Sarma, M. M., & Mohapatra, D. (2018). Test impact and test design: Insights from the Syrian National Baccalaureate Examination of English. *The ELT Practitioner*. 5 (2). Retrieved December, 3, 2019 from <https://shamra.sy/academia/show/5b9eb9dd53fda>
- Moss, S. (2014). *Fit Indices for Structural Equation Modeling*. Retrieved July, 8, 2020 from <https://www.sicotests.com/psyarticle.asp?id=277>.
- Nimehchisalema, V., & Bhattib, N. (2019). A review of literature on Language Assessment Literacy in the last two decades (1999-2018). *International Journal of Innovation, Creativity and Change*. 8(11). 44-59.
- Popham, J. W. (2004). Why assessment illiteracy is professional suicide. *Educational Leadership*, 62 (1), 82.
- Quaigrain, K., & Arhin, A. K. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation, *Cogent Education*, 4(1), 1-11.
- Riazi, A. M., & Candlin, N.C. (2014). Mixed-methods research in language teaching and learning: Opportunities, issues and challenges. *Language Teaching*, 47(2), 135-173.
- 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.
- Stiggins, R. (2014). Improve assessment literacy outside of schools too. *Phi Delta Kappa*, 96(2), 67-72.
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappa*, 72(7), 534-39.



- Syrian Ministry of Education and scientific research (n.d.). Retrieved December 2, 2019, from <http://moed.gov.sy/site/>.
- Tavassoli, K., & Farhady, H. (2018). Assessment knowledge needs of EFL teachers. *Teaching English Language, 12*(2), 45-65.
- Tomé-Fernández, M., Fernández-Leyva, C., & Olmedo-Moreno, E. M. (2020). Exploratory and Confirmatory Factor Analysis of the Social Skills Scale for Young Immigrants. *Sustainability, 12*(17), 2-20.
- 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.
- Vogt, K., Tsagari, D., & Csépes, I. (2020). Linking learners' perspectives on language assessment practices to teachers' assessment literacy enhancement (TALE): Insights from four European countries. *Language Assessment Quarterly, 1*-24.
- Watmani, R., Asadollahfam, H., & Bahram 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.
- Willis, J., Adie, L., & Klenowski, V. (2013) Conceptualizing teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher, 40*(2), 241–256.
- Yan, X., & Fan, J. (2020). "Am I qualified to be a language tester?": Understanding the development of language assessment literacy across three stakeholder groups. *Language Testing 0*(0). 1-28.

## Appendix

### *Adaptation of Fulcher's (2012) LAL Survey*

1. What assessment tips and skills do you think English language teachers ought to be familiar with?
2. Please look at each of the following topics of language testing.

For each one, please specify your level of familiarity with these testing topics:

Indicate your response as follows:

1= not at all proficient

2= not very proficient

3= fairly proficient

4= proficient

5= highly proficient

A. History of language testing	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
B. Procedures in language test design	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
C. Deciding what to test	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
D. Writing test specifications/blueprints	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
E. Writing test tasks and items	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
F. Evaluating language tests	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
G. Interpreting scores	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
H. Test analysis	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
I. Selecting tests for your own use	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
J. Reliability	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
K. Validation	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
L. Use of statistics	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
M. Rating performance tests (speaking/writing)	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
N. Scoring closed-response items	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
O. Classroom assessment	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
P. Large-scale testing national test 9 <sup>th</sup> , 12 <sup>th</sup>	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
Q. Standard setting	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
R. Preparing learners to take tests	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
S. Side effects of test on teaching	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
T. Test administration	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
U. Ethical considerations in testing	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
V. The uses of tests in society	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5
W. Principles of educational measurement	1 ◦ 2 ◦ 3 ◦ 4 ◦ 5

3. Have you ever read or used a language testing book to write English exams?

- Yes
- No

4. What do you think are essential topics in a book on practical language testing?

5. How would you rate your knowledge and understanding of language testing?

1 = very poor

2 = poor

3 = average

4 = good

5 = very good

6. Are you male or female?

- Female | ◦ Male

7. What is your age range?

◦ 21 - 30

◦ 31 - 40

◦ 41 - 50

◦ 51 - 60

◦ 61 - 70

8. Please select your current educational level

- High School Graduate
- BA degree
- MA degree
- Doctorate
- Other

9. Which Syrian university did you take your degree from?

- Damascus University
- University of Aleppo
- Tishreen University
- Al-Baath University
- Al-Furat University
- University of Hama

10. Please specify your experience in teaching English:

- 0-5 years
- 5-10 years
- 10-20 years
- More than 20 years

11. Where do you teach English?

- Institute
- University
- School

