

Developing and Validating a Questionnaire to Assess EFL Teachers' Classroom-Based Assessment Literacy

Elham Banisaeed 

Ph.D. Candidate of TEFL, Karaj Branch, Islamic Azad University, Karaj, Iran

Mohammad Hashamdar* 

Assistant Professor of TEFL, Karaj Branch, Islamic Azad University, Karaj, Iran

Kobra Tavassoli 

Assistant Professor of TEFL, Karaj Branch, Islamic Azad University, Karaj, Iran

Received: January 13, 2024; **Accepted:** June 30, 2024

Abstract

Classroom-based assessment (CBA) as one of the constructs of formative assessment has been considered highly significant in recent years. Consequently, various tools have been designed to investigate teachers' CBA needs and deficiencies ignoring different levels of teachers' CBA literacy. Thus, the present study researchers developed and validated a classroom-based assessment literacy questionnaire (CALQ) to determine teachers' levels of CBAL. To do so, an inclusive review of the literature was accomplished to retrieve major themes and components of CBAL, and then a series of interviews were conducted with five assessment experts and 13 experienced EFL teachers in accordance with Pill and Harding's (2013) Model of LAL, Hill and McNamara's (2012) scope and dimensions of CBA in addition to teachers' assessment literacy beliefs. Accordingly, a questionnaire (CALQ) including 41 items was developed. To inquire the reliability and validity of the CALQ, 318 EFL teachers were selected through non-probability convenience sampling and asked to answer the questionnaire. The outcomes of the Cronbach's alpha demonstrated a proper reliability index, and factor analysis products clarified that items loaded on six factors named as illiteracy (6 items); nominal literacy (11 items); functional literacy (6 items); procedural and conceptual literacy (6 items); multidimensional literacy (6 items); and assessment literacy beliefs (6 items). Besides, CALQ is considered advantageous in assessing teachers' CBAL and facilitating materials preparation to design instructional courses and develop EFL teachers' CBAL, based on the conclusions of structural equation modeling (SEM), which proved that the Model enjoyed good psychometric features.

Keywords: Assessment Literacy (AL), Language Assessment Literacy (LAL), Classroom-based Assessment Literacy (CBAL), Questionnaire, Validation

*Corresponding author's email: mohamad.hashamdar@kiaiu.ac.ir

INTRODUCTION

Assessment has always been regarded essential in the area of language instruction as long as it is the only tool to ascertain teachers whether students have achieved the determined goals. Additionally, according to Turner (2012 cited in Fulcher & Davidson 2012) since Black and William's (1998) influential paper on classroom formative assessment has been published, an increasing interest has been identified in classroom-based assessment (CBA) and its capability for increasing learning. The main reason relies on the fact that the information required for teachers to succeed in evaluating learning cannot be achieved by traditional ways of testing (e.g., multiple choice, essay, and reading tests).

On the other hand, along with shifts in teaching from method to post-method, testing also witnessed shifts from high-stakes standard tests to local and teacher-made tests to increase learning. These alterations highlighted the importance of CBA (Farhady, 2019).

Even though formative assessment and consequently CBA have been identified as beneficial types of assessment (Dehqan & Asadian Sorkhi, 2020) no comprehensive classroom-based assessment literacy (CBAL) questionnaire has been developed to investigate the CBA knowledge of language teachers who are evaluating students in classroom contexts and consequently, it is highly demanded to identify the extent teachers are aware of principles of CBA. Thus, the present study concentrates on the purpose that a standard scale which classifies teachers according to their knowledge of CBAL could play a vital role in investigating the concept of CBA, and accordingly, leads to teachers professional development which is considered highly influential in improving teachers' practices (Khany & Azimi Amoli, 2016). Also, the study seeks to find the factors influencing teachers' CBAL in their view points. The reason for this research is expanding desire to regard classroom teachers as the assessors and also increasing cognizance of the effect of assessment on learning. Moreover, such research is significantly required to enhance the education outcomes related to students' achievements

being assessed by the teachers in the classroom context.

LITERATURE REVIEW

Assessment literacy (AL) was first suggested by Stiggins (1991). Since then, the concept of AL has been discussed and investigated by different studies. The first attempt encompassing research in language assessment literacy (LAL) proposed the 1990 “Standards for Teacher Competence in Educational Assessment of Students”. These Standards provided the field with the required knowledge to conduct further research. These attempts resulted in different studies investigating LAL components and knowledge bases while offering various models.

LAL was viewed as including different components in the first group of models (Davies, 2008; Fulcher, 2012; Taylor, 2013; Xu & Brown, 2016) while it was regarded as composed of different dimensions and levels in the second one including Pill and Harding’s (2013) Model of LAL which was developed based on science literacy and mathematics (Coombe et al., 2020) and classified language teachers’ literacy into different levels, from illiteracy to multidimensional literacy.

Along with different proposed models of LAL, the literature includes a variety of studies all attempting to present different measures of LAL. Most of the studies have included quantitative scales mainly derived from the 1990 Standards as the underlying framework. Generally, the proposed measures including Assessment Literacy Inventory (ALI) (Campbell, 2002), Classroom Assessment Literacy Inventory (CALI) (Mertler, 2003) and Teacher Assessment Literacy Questionnaire (TALQ) (Plake et al., 1993) presented some content-based items to the respondents to provide their answers. Later, Campbell and Mertler (2005) provided respondents with some scenario-based items and examined their answers. As it could be understood, the all 8 measures of LAL for EFL teachers were based on the 1990 Standards and failed to explore the more recent dimensions of LAL (e.g., social) and mostly divided teachers into two groups of assessment illiterate and literate ones ignoring the fact that LAL as a continuum includes different levels. The

above-mentioned reasons, highlighted the necessity of developing a measure of CBAL for EFL teachers which considers all dimensions of LAL and also provides a scale to identify different levels of EFL teachers' assessment knowledge to facilitate assessment course design and preparation for professional development purposes.

McNamara (2001 cited in Hill & McNamara, 2012) defined classroom-based assessment as any thoughtful, prolonged, and definite observation (by teachers or students) to identify characteristics of students' performance and employ the information as assistance to articulate learning objectives. Hill and McNamara (2012) who conducted empirically based research, presented a comprehensive framework to investigate CBA. The framework suggested three dimensions named "evidence", "interpretation", and "use". The first dimension, "evidence", deals with the kind of data collected by the teacher, the approach of assessment undertaken by the teacher, and the role played by the teacher and the students. The second dimension, "interpretation", concentrates on reflection and the criteria for assessment used by the teacher. The third one, "use", focuses on the purpose and agent of assessment.

Consequently, Chappuis et al., (2012 cited in Yamtim & Wongwanich, 2014) delineated CBAL as the required information and skills to collect data about learners' attainment and efficiently employ the process of assessment and results to increase the potency of teachers' instruction and learners' attainment. Tsagari (2016) mentioned that the ignorance of CBAL training impedes teachers' innovative use of assessment techniques. Although CBA has been emphasized in recent years, research has proved that teachers are to some extent unprepared to administer this kind of assessment(e.g., Crusan et al., 2016) and a study executed by Narathakoon et al., (2020) demonstrated that teachers often employed final and mid-term examinations in addition to student observation as different tools for classroom assessment.

Since CBA is directly related to the active role played by teachers in the classroom, EFL teachers' beliefs regarding assessment become prominent. The concept of teachers' beliefs has always been challenging to be defined. Borg (2003) mentioned that teachers' beliefs encompass a wide range of

knowledge and assumptions regarding theory and practice.

PURPOSE OF THE STUDY

The current work strives to scrutinize the underlying constituents of teachers' CBAL in terms of knowledge, skills, and beliefs. It also tries to adapt Pill and Harding's (2013) Model of LAL in alignment with CBA scope, dimension, and teachers' classroom assessment beliefs. The study is also considered novel since it focuses not only on developing a questionnaire to assess EFL teachers' CBAL but also on conceptualizing CBAL as a construct including six components of illiteracy, nominal literacy, functional literacy, procedural and conceptual literacy, multidimensional literacy, and assessment literacy beliefs. The subsequent research questions were proposed to attain this goal:

1. What are the fundamental constituents of the classroom-based assessment literacy questionnaire(CALQ)?
2. What are the psychometric features of the classroom-based assessment literacy questionnaire(CALQ)?
3. To what extent does the structural model of classroom-based assessment literacy questionnaire (CALQ) fit the hypothetical model formed by consistent literature?

METHOD

Participants

The participants in the qualitative stage included five assessment experts and 13 experienced teachers. The participating experts were all male, aged from 43 to 60, and Ph.D. holders in TEFL. They have been teaching in different universities for 18 to 35 years. They all have published at least two papers related to assessment in prestigious journals.

Also, 13 EFL teachers experienced in teaching English for more than ten years participated in the inquiry. The researchers attempted to choose those who reasonably regarded skillful in teaching and testing to be specifically able to provide comprehensive answers. The teachers were all females, aged

from 29 to 51 years. They had BA and MA in TEFL, Translation, and Literature. Also, they have been teaching English in schools and institutes in different cities. The participants were selected using nonprobability convenience sampling (Best & Kahn, 2006).

The newly developed questionnaire consisted of 50 items, so 50 volunteer EFL teachers asked to participate in the piloting phase. Participants included 37 female and 13 male teachers teaching English in various institutes and schools in Iran. Considering their teaching experience, 24% were experienced less than 5 years, 32% were experienced 6 to 10 years, and 44% were experienced more than 10 years. Regarding educational degrees, 42% had a BA, and 58% had an MA. 86% of the participants majored in English and 14% in non-English disciplines.

In the administration phase, the questionnaire was filled out and returned by 342 teachers who were chosen through non-probability convenience sampling, out of which, 24 dropped out. The omitted respondents either left the majority of the items blank or selected similar choices for the whole or significant parts of the statements. Thus, the quantitative phase was conducted with 318 teachers. The following table presents the participants' demographic information.

Table 1: Participants' Characteristics in the Administration Phase

Participants Characteristics		Frequency	Percentage
Gender	Male	153	48%
	Female	165	52%
Years of Experience	Less than 5 years	119	37%
	6 to 10 years	107	33%
	More than 10 years	92	30%
Educational Degree	Diploma	10	3%
	Associate Degree	57	19%
	BA	179	56%
	MA	72	22%
Major	English	235	74%
	Non-English	83	26%
Total		318	100%

They included 153 male and 165 female EFL teachers with different years of experience, including 119 teachers with less than five years of experience, 107 experienced 6 to 10 years, and 92 experienced more than ten years. The participants had various educational degrees, only 10 had diplomas, 57 had associate degrees, 179 BA, and 72 had MA. Regarding their major, 235 teachers majored in English, while 83 majored in non-English disciplines. They participated in the study from different cities, teaching in schools and institutes nationwide. The participants were selected according to the non-probability sampling. All the participants in the qualitative phase (interview) and the quantitative phase (piloting and administration phase of the questionnaire) assured of the confidentiality of the data they provided the researchers with and informed of the fact that codes were used instead of their names (e.g., A, B, C).

Pallant's (2016) strategy regarding sample size estimate used in the current study, which asserted that for each item, 5 participants would be a sufficient sample proportion to choose the most suitable number of people to answer the newly developed questionnaire. Therefore, a minimum sample size of 250 participants was needed for the current study since the CALQ includes 50 items.

Instrumentation

The present study employed a variety of instruments, including several semi-structured interviews with experienced EFL teachers and assessment experts, and a Likert-scale-based questionnaire to collect the required data in different phases.

The present study utilized three models as instruments to collect data. The first one was Pill and Harding (2013) Model of LAL which assumed LAL as a sequence including five ranks named illiteracy, nominal literacy, functional literacy, procedural and conceptual literacy, and multidimensional literacy. The second one was the CBA dimensions suggested by Hill and McNamara (2012) which seemed comprehensive to investigate the concept.

Thirdly, as long as CBA is carried out in a classroom context by teachers, teachers' beliefs regarding assessment as one of the components of Borg's (2003) concept of "teacher cognition" were employed to develop interview questions.

A semi-structured interview including eight questions (Appendix A) was administered in English after conducting a comprehensive literature review. All questions were based on the five components of Pill and Harding's Model of LAL, three dimensions of CBA proposed by Hill and McNamara (2012), and teachers' CBA beliefs. The interviewer asked the interviewees to expand their answers in detail to obtain rich data for further investigation. Furthermore, the researchers developed a valid and reliable five-point Likert-scale-based questionnaire composed of 41 items that dealt with features of teachers' CBAL (Appendix B). The details of its development are mentioned in the procedure section.

Data Collection Procedure

Since the current study is an exploratory sequential mixed-methods research (Creswell & Creswell, 2018), the demanded data were collected in two phases: qualitative and quantitative. The qualitative stage included an inclusive review of literature on AL, LAL, and CBAL to identify the conceptual framework based on which the eight questions for the semi-structured interview were developed. A group of five experts in the field of assessment closely examined the questions in terms of contents and wording appropriateness to prove the credibility of the interview questions. Subsequently, a semi-structured interview was administered with five assessment experts and 13 experienced EFL teachers. Each interview lasted 30 to 40 minutes, and was audio-recorded, transcribed, and coded according to Merriam and Tisdell framework (2016). The retrieved themes were cross-checked with the literature. Then, considering the extracted ones, a questionnaire consisting of 50 items (CALQ) was developed. Table ۲ presents the introductory elements and topics in CALQ:

Table ۲: Introductory Elements and Topics in CALQ

LAL Components	CBA Dimensions	Interview Questions	Themes	Example of CALQ Items
Illiteracy (Ignorance of language assessment concepts and methods)	Evidence, Interpretation, Use	In your opinion, what kinds of teachers are considered assessment illiterate?	1 not knowing assessment concepts and method 2 ignoring the importance of contextual factors 3 not being able to put assessment knowledge into practice 4 not employing multiple types of assessment and alternative assessment 5 lacking creativity and need analysis 6 not aligning assessment with learning goals	Assessment illiterate teachers are those who are <u>not</u> able to apply their knowledge, construct a test, and develop test items.
Nominal Literacy (Recognizing that a peculiar word pertains to assessment, but may show a misunderstanding)	Evidence, Interpretation, Use	Since it was believed that assessment experts and experienced EFL teachers are nominally assessment literate, this component	A list of assessment key terms and concepts, including authentic assessment, formative vs. summative assessment, cut-off score, and qualitative vs. quantitative	I am familiar with authentic assessment.

		was not included in the interview.	assessment ... was developed.	
Functional Literacy (reliable comprehension of primary assessment words and topics)	Evidence, Interpretation, Use	What is Validity/ Reliability/ Formative Assessment/ Summative Assessment?	The interviewees were asked to define some assessment key terms and concepts. CALQ, included six items, some of which described the key words in the wrong way to examine the respondents' sound understanding of the assessment key words.	Authentic assessment is a type of assessment in which students are required to implement pedagogical tasks that illustrate meaningful employment of necessary knowledge and skills.
Procedural and Conceptual Literacy (comprehending basic terms of the field and employing knowledge)	Evidence, Interpretation, Use	What are the different steps in constructing language tests?	The interviewees were asked to elaborate on the major steps of test construction: Identifying the purpose and the form of the test Preparing the items Reviewing the items Pretesting the items	Identifying the aim and the form of the test is among the major steps of test construction.
Multidimensional Literacy (awareness ranging above typical themes, including philosophical, historical, and social aspects of assessment)	Evidence, Interpretation, Use	What is the signification of assessment in education? What is the relationship between assessment and society?	Assessment is critical in education. Assessment plays the role of a criterion in society to select more competent applicants in different fields.	Without assessment, education can achieve predetermined goals.

	Since when was the assessment considered vital?	Assessment has always been vital.	
Assessment Literacy Beliefs	Do you believe in the relationship between experience, education, and assessment literacy?	Majoring in English and teaching and testing experience lead to assessment literacy.	Assessment illiterate teachers are those who are <u>not</u> educated in the English major.
	What is the purpose of assessment?	The fundamental function of assessment is improvement and development.	The partial purpose of assessment is improvement and development.

As Table 3 illustrates, the first component was illiteracy, and teachers and experts were asked to elaborate on the characteristics of the assessment illiterate teachers. They listed some features, such as not knowing assessment concepts and methods, as the main characteristics of assessment illiterate teachers. Since assessment illiterate teachers have been defined in the Model as being unaware of language assessment methods and concepts, the mentioned response by the interviewees was included as one of the items of the illiteracy component in CALQ (i.e., assessment illiterate teachers are those who do not know assessment concepts and methods).

The second component in the Model was nominal literacy, and since assessment experts and experienced teachers were believed to be familiar with assessment key terms and concepts, no question regarding nominal literacy was asked in the interview. Items targeted to measure the participants' familiarity with assessment key terms and concepts were only included in CALQ (e.g., I am familiar with assessment).

The next question was related to functional literacy, which was the third component of the Model. The interviewees were required to define some critical terms of assessment, such as validity, reliability, and summative vs. formative assessment. The questions aimed to examine experts' and teachers' sound understanding of the terms. Later, the terms were defined incorrectly and included among items of CALQ to measure respondents' proper knowledge of the terms (e.g., Formative assessment is operated at the termination of the lesson while summative assessment is undertaken during the learning process).

The fourth component, which was called procedural and conceptual literacy, included questions regarding the significant steps of test construction. Experts and teachers named some major stages of test construction. Next, the researchers used these to develop CALQ items (e.g., identifying the aim and the form of the test is among the significant stages of test construction).

The fifth component in Pill and Harding's Model of LAL was multidimensional literacy, consisting of historical, philosophical, and social dimensions. To evaluate the interviewees' philosophical dimension of assessment, they were asked to discuss the philosophy of assessment, and the mentioned concepts were used to develop CALQ items (e.g., without assessment, education can achieve predetermined goals). The social dimension was investigated by asking interviewees to elaborate on the relationship between assessment and society, and their provided answers were employed to write CALQ items (e.g., assessment plays the role of a criterion in society to select more competent applicants in different fields). The historical dimension was explored by asking interviewees to discuss the first-time that assessment was considered vital. The provided answers were used to write CALQ items (e.g., assessment has always been vital).

The last component was assessment literacy beliefs, which were investigated by asking questions such as what the purpose of assessment is and whether they believe in the relationship between experience, education, and assessment literacy. CALQ items (e.g., assessment illiterate teachers are

those not educated in the English major and the partial purpose of assessment is improvement and development) were among items considering assessment literacy beliefs. Additionally, as Table 2 indicated, all interview questions were related to CBA scope and dimensions.

Moreover, the researchers developed valid and reliable five-point Likert-scale-based interview results constituting 50 items, including components of teachers' CBAL. CALQ, included six components named illiteracy (7 items), nominal literacy (11 items), functional literacy (8 items), procedural and conceptual literacy (11 items), multidimensional literacy (7 items), and assessment literacy beliefs (6 items). A "five-point Likert scale" was employed for the study (*1= strongly disagree, 2= disagree, 3= neither agree nor disagree, 4= agree, 5= strongly agree*). There were also reversed items (items 21, 23, 26, 42, 43, 50) for which the "five-point Likert scale" valued the opposite.

The newly developed questionnaire was piloted with 50 EFL teachers using non-probability convenience sampling. All CALQ items were replied to by all the respondents. Considering the COVID-19 outbreak, an online questionnaire was designed employing the Google Forms platform and the participants were asked to respond to it. Later, a Cronbach's alpha was run to determine and remove questionable items (Dörnyei, 2003), and exploratory factor analysis (Riazi, 2016) was run to identify the primary components of the CALQ. Then, following a similar methodology, the final draft of the CALQ was distributed to 318 EFL teachers from different schools and institutes in Iran chosen through non-probability convenience sampling.

Data Analysis

Ultimately, the data achieved from the main participants were investigated through the following statistical analyses. Using the IBM SPSS software (version 26), the newly designed CALQ was subjected to Cronbach's alpha and exploratory factor analysis (EFA) to investigate its reliability and determine the underlying components of the 50 items of the instrument. Conforming to Riazi (2016), whereas EFA is employed as a statistical test to

identify the primary constructs of a concept by compacting the data to a more controlled number of variables, confirmatory factor analysis (CFA) is a statistical test employed to approve the component design of a group of observed variables. Thus, in the present study, CFA through structural equation modeling (SEM) was run using IBM AMOS 22 to inquire and ensure the fit of the CALQ model, including six measurement models.

RESULTS

CALQ was piloted with 50 EFL teachers to calculate its reliability before being answered by the primary sample in an authentic context. In this piloting phase, the overall questionnaire enjoyed a reliability index of .870. The reliability indices, for the components were as follows: Illiteracy ($\alpha = .734$), Nominal Literacy ($\alpha = .947$), Functional Literacy ($\alpha = .825$), Procedural and Conceptual Literacy ($\alpha = .700$), Multidimensional Literacy ($\alpha = .752$), and Assessment Literacy Beliefs ($\alpha = .862$) which was a sign of an appropriate reliability index according to George and Mallery (2020).

The data obtained from 318 EFL teachers in the administration process was first checked for any substantial univariate and multivariate outliers. The assumption of univariate outliers was tested by calculating the standardized scores (z scores) for particular items of the CALQ. None of the statistics were higher than ± 3.29 ; thus, it was determined that the present data did not suffer from any considerable univariate outliers (Table 1, Appendix C). It should be noted that the criteria of ± 3.29 is suggested by Tabachnick and Fidell (2014, p 107). The assumption of lack of any significant multivariate outliers was tested by calculating the Mahalanobis Distances (MD). As mentioned by Tabachnick and Fidell (2014) and Watkins (2021), the MD indices should be evaluated against the decisive value of chi-square at .001 levels for 50 items of the CALQ, i.e., 88.66. The maximum MD value of 72.20 was lower than the decisive value of 88.66. Thus, it was identified that the assumption of lack of multivariate outliers was also assumed (Table 2, Appendix C). The univariate normality of the data was examined through skewness and kurtosis indices. Since the values of skewness and kurtosis were within the limits of

± 2 (Bae & Bachman, 2010; and George & Mallery, 2020), it was determined that the premise of univariate normality was met. The multivariate normality of the data was investigated by Mardia's index. The Mardia's index should be measured against the criteria of ± 3 (Bae & Bachman, 2010; Zhu et al., 2019). The results indicated that the premise of multivariate normality was also assumed (Table 3, Appendix C).

Table 3 shows Cronbach's alpha reliability indices for the overall CALQ and its six components. The whole questionnaire enjoyed a reliability index of .876. The reliability indices for the factors were as follows: Illiteracy ($\alpha = .850$), Nominal Literacy ($\alpha = .930$), Functional Literacy ($\alpha = .870$), Procedural and Conceptual Literacy ($\alpha = .889$), Multidimensional Literacy ($\alpha = .868$), and Assessment Literacy Beliefs ($\alpha = .868$). The reliability indices mentioned above can be considered appropriate, as noted by Fryer et al., (2018), and Harrison et al., (2021), who asserted that Cronbach's alpha value of .70 is the sufficient reliability index for a questionnaire.

Table 3: Reliability Statistics

	Cronbach's Alpha	N of Items
Illiteracy	.850	6
Nominal Literacy	.930	11
Functional Literacy	.870	6
Procedural and Conceptual Literacy	.889	6
Multidimensional Literacy	.868	6
Beliefs about Assessment Literacy	.868	6
Total	.876	41

Exploratory Factor Analysis (EFA) employing the principal axis factoring method and varimax rotation was run to examine the primary components of the CALQ after dropping out the nine items that did not load under their related components to probe the construct validity of CALQ. Before discussing the results, the rotation method and number of factors extracted should be justified.

The second EFA was run using the varimax rotation method since correlations among the constructs (Table 4, Appendix C) were not all higher than $\pm .32$ (Dagdag et al., 2020). In other words, there were no significant

correlations among all constructs. It should be mentioned that varimax rotation, a member of orthogonal methods, assumes that factors are not correlated.

Two types of parallel analyses, i.e., computational and graphical, were employed to identify the optimum number of components extracted. Watkins' Parallel Analysis compares the initial eigenvalues, the total percentage of variance explained by an item, against the simulated ones. The factors whose initial eigenvalues are higher than the simulated ones are retained. The results of Watkins' Parallel Analysis (Table 5, Appendix C) suggested six components extracted as the primary factors of the CALQ.

Eventually, Revelle (2020) also developed a graphical method through the R Package "psych," similar to scree plots produced by SPSS, to decide how many factors should be extracted. This method compares the initial eigenvalues against the bootstrapped ones and shows the number of factors removed. Similarly, this method also suggested six factors as primary constructs of CALQ.

The KMO index of .937 indicated that the current sample size was "marvelous" according to Field's (2018) classification of KMO indices for running EFA (Table 6, Appendix C). The significant results of the sphericity test ($\chi^2(820) = 6275.99, p < .05$) demonstrated that the correlation matrix used to run EFA was factorable. The EFA identified six components as the primary factors of the 41 items of the CALQ (Table 7, Appendix C), which counted for 53.42 percent of the total variance. The 41 items loaded under the respective factor loadings are as follows:

Table 4: Components of Classroom-Based Assessment Literacy Questionnaire

Components	N of Items	Items
Illiteracy	6	3-4-5--8-9- 10
Nominal	11	11 to 20, 31
Functional	6	21-22-23-24-26-27
Procedural and Conceptual	6	29-30-32-33-36-37
Multidimensional	6	6-41-42-46-49- 50
Beliefs	6	1-2-43-45-47- 48

A confirmatory factor analysis (CFA) was run using IBM AMOS 22 to investigate the fit of the CALQ model. The Model comprises six measurement models (Appendix C, Figure 1), whose fit was discussed before discussing the overall Model. Figure 2 represents the final model of CALQ.

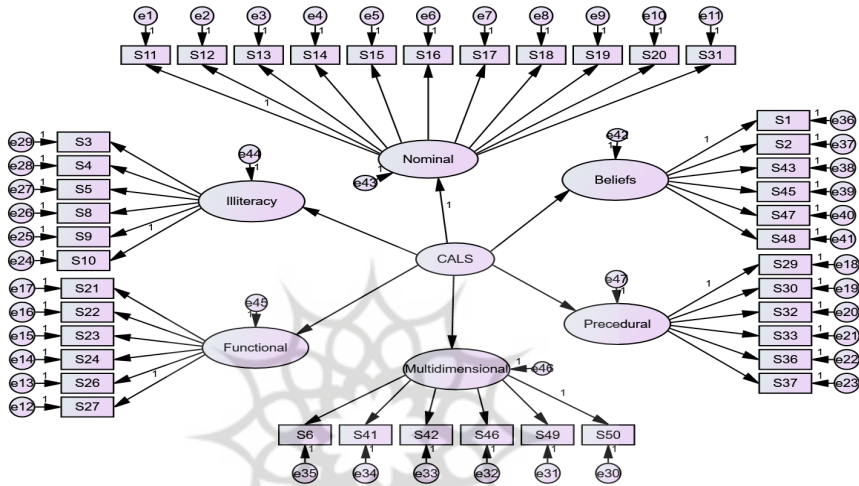


Figure 2: Main Structural Equation Model

The Model enjoyed a good fit as presented in Table 8, Appendix C. The non-significant chi-square index of the badness of fit advocated the fit of the primary CALQ Model ($\chi^2(773) = 450.78, p = 1.00$). The ratio of chi-square over the degree of freedom, i.e., .583, was lower than 3. The SRMR index of .024 was lower than .05. The RMSEA of .000, and its confidence intervals [.000, .000] were lower than .05. All these indices maintained the fit of the SEM measurement model. Table 5 displays all fit indices for the CALQ and its six components:

Table 5: All Fit Indices for Six Measurement Models

	Indices	“Illiteracy”	“Nominal”	“Functional”	“Procedural/Conceptual”	“Multidimensional”	“Beliefs”	Criteria	Fit
Absolute	X ²	5.43	57.40	3.93	5.97	14.43	12.10	---	---
	df	9	44	9	9	9	9	---	---
	p	.795	.085	.916	.742	.108	.208	>.05	Good Fit
	X ² Ratio	.604	1.30	.437	.664	1.60	1.34	<3	Good Fit
	SRMR	.014	.022	.011	.013	.021	.020	<.10	Good Fit
	RMSEA	.000	.031	.000	.000	.044	.033	<.05	Good Fit
	CI	[.000,.041]	[.000,.052]	[.000,.023]	[.000,.046]	[.000,.084]	[.000,.076]	<=.10	Good Fit
	PCL	.975	.933	.993	.965	.550	.694	>.05	Good Fit
	OSE	.994	.970	.996	.994	.985	.988	>=.90	Good Fit
	GFI	.986	.964	.992	.989	.969	.974	>=.90	Good Fit
Incremental	RFI	.986	.964	.992	.989	.969	.974	>=.90	Good Fit
	TLI	1	.991	1	1	.988	.993	>=.90	Good Fit
	CFI	1	.993	1	1	.993	.996	>=.90	Good Fit
	NFI	.992	.971	.995	.994	.982	.984	>=.90	Good Fit
	IFI	1	.993	1	1	.993	.996	>=.90	Good Fit
Hoelter (Sampling Adequacy)	988	334	1365	897	372	444	>20	Adequate	

DISCUSSION

Whereas there are not sufficient studies in the area of LAL surveys, which is the central space that exists in the literature and causes researchers to turn to such surveys, the current study purposed to explore the underlying

components of CBAL and accordingly, design and validate a questionnaire to assess teachers' CBAL. The results proved the reliability and validity of the currently designed CALQ while 41 items loaded on six factors: 1) illiteracy, 2) nominal literacy, 3) functional literacy, 4) procedural and conceptual literacy, 5) multidimensional literacy, and 6) assessment literacy beliefs. Moreover, SEM outcomes demonstrated that the Model enjoyed good psychometric features.

The first component in CALQ, illiteracy, includes six items to investigate the characteristics of CBA illiterate teachers. In accordance with the current study findings, it could be argued that EFL teachers need to possess a comprehensive width of understanding regarding concepts and methods of CBA, including knowledge of assessment types, assessment theories, contextual factors, practical knowledge, multiple types of assessment, and assessment alignment with learning goals as the contributing factors to CBAL lack of which leads to CBA illiteracy. The present study findings in this regard are supported by Rahimi et al., (2021), who investigated the main constituents of LAL in Iran and proposed a model based on three components, namely "awareness of language pedagogy," "assessment principles and interpretation," and "assessment policy and local practices". Similarly, Khodashenas et al., (2022) conducted a study in which an inventory called "Teachers Assessment Literacy Needs (TALNs)" was developed, demonstrating that teachers' knowledge of assessment processes and consequences was viewed as one of the major requirements of their CBAL.

The second factor in CALQ is nominal literacy. Eleven items in CALQ, targeted to measure teachers' familiarity with assessment key terms and concepts. Considering the investigation of EFL teachers' level of familiarity with assessment key terms and concepts, Sasmaz-Ören and Ormancı (2011) accomplished a survey study to examine teacher candidates' familiarity with alternative assessment. Similarly, Farikhah et al., (2022), in a case study, explored a novice teacher's acquaintance with "assessment for learning" in language classrooms.

The third factor in CALQ is functional literacy. All six items of CALQ

regarding this component evaluate the teachers' sound understanding of significant assessment key terms. CALQ focuses on concepts such as assessment, alternative assessment, authentic assessment, criterion-referenced vs. norm-referenced assessment, formative vs. summative assessment, and qualitative vs. quantitative measurement, which are key terms related to the evidence and interpretation of CBA dimension according to Hill and McNamara (2012). In a similar study, Nikmard and Mohamadi (2020) attempted to develop a questionnaire to assess ELTs' assessment literacy. They proposed the instrument by undertaking a series of steps, including a comprehensive review of the literature and identifying four components for teachers' AL: "validity", "reliability", "interpretability of the results", and "efficiency". After interviewing participants and the piloting step, they developed and validated a questionnaire for assessing ELTs' assessment literacy consisting of 25 items employing a five-point Likert scale.

The fourth factor in CALQ was called procedural and conceptual literacy. Consequently, CALQ includes six items aiming to assess not only teachers' conceptual knowledge of assessment key terms and concepts such as validation and standardized testing but also teachers' procedural knowledge of assessment, including significant test steps construction. Aria et al., (2021) implemented a study aiming to explore Indonesian EFL secondary teachers' attitudes of classroom-based assessment practice. They developed a survey including four significant components of "planning assessment principles," "assessment implementation principles," "monitoring assessment principles," and "disseminating assessment principles". Similarly, to investigate EFL teachers' assessment literacy, Rastegar et al., (2022) developed a questionnaire identifying nine significant components of assessment literacy in the Iranian context, two of which are naming "test construction", "recognizing test type, distinction, and function".

The fifth factor in CALQ is multidimensional literacy. CALQ includes six items to explore teachers' opinions regarding philosophical, historical, and social aspects of assessment. Among the three dimensions, the only one

investigated by the previous studies is the social dimension. Yan and Pastore (2022) in an attempt to develop and validate the “teacher formative assessment literacy scale (TFALS)”, designed a tool based on a “three-dimensional formative assessment model”, including “conceptual”, “practical”, and “socio-emotional” components. The instrument consisted of 7 items targeting to highlight the importance of socio-emotional considerations of supplying students with assessment feedback. Similarly, Tajeddin et al., (2022) conducted a study to assess teachers’ grasped classroom-based assessment awareness and practice via a “classroom-based language assessment literacy” scale. They proposed a model for CBA based on four components of “assessment purpose and grading,” “assessment ethics,” “student involvement,” and “feedback and assessment interpretation and communication”.

Finally, the last factor in CALQ is devoted to the assessment literacy beliefs. Since teachers’ practice in CBA is highly guided and affected by their beliefs and attitudes regarding CBA (Alonzo et al., 2021; Barnes et al., 2015; Crusan et al., 2016; Dashti, 2019; Munoz et al., 2012; Toth & Csapo, 2022; Unal & Unal, 2019), any attempt to investigate teachers’ CBAL without considering their assessment literacy beliefs lacks comprehensiveness. On the other hand, Borg (2003) introduced “teacher cognition” as the combination of different interrelated factors affecting teachers’ CBA performance, including teachers’ beliefs, knowledge, skills, and conceptions. In consequence, the current study researchers decided to include teachers’ assessment literacy beliefs as one of the components of CALQ. The questionnaire contains six items aiming to explore teachers’ assessment literacy beliefs, such as how teachers’ majors and experience affect their practice of assessment and also, attitudes toward the purpose of assessment. Adopting a mixed-methods approach, Alyami (2022) developed a questionnaire to explore teachers’ beliefs concerning classroom assessment in the Saudi background. She reported the classification of teachers’ beliefs as follows: “general beliefs about CBA”, the “purposes of assessment”, “assessment methods” and “the role of students in assessment”, “aligning

assessment and learning objectives”, “frequency of assessment”, and “provision of feedback”.

The present study could be regarded as innovative in different respects. CALQ compared with previous studies, not only classifies EFL teachers’ knowledge based on their CBAL level but also deals with teachers’ beliefs regarding CBAL, which has been neglected by former studies.

CONCLUSION AND IMPLICATIONS

The priority of CBAL has been stressed by various investigations (e.g., Fitriyah et. al., 2022; Kingston & Nash, 2011; Tofighi & Ahmadi Safa, 2023). Accordingly, it has been emphasized to develop instruments to measure teachers’ CBAL (e.g., Gotch & French, 2014). Therefore, the undertaken study was conducted to design and certify a classroom-based assessment literacy questionnaire (CALQ) to assess teachers’ CBAL. CALQ includes six components (i.e., illiteracy 6 items, nominal literacy 11 items, functional literacy 6 items, procedural and conceptual literacy 6 items, multidimensional literacy 6 items, and assessment literacy beliefs 6 items). The questionnaire is designed based on Pill and Harding’s (2013) Model of LAL, Hill and McNamara’s (2012) scope and dimensions of CBA, in addition to Borg’s (2003) model of “teacher cognition”. Applying exploratory factor analysis, confirmatory factor analysis, and structural equation modeling as analytical procedures, CALQ proved to be valid and reliable and enjoyed good psychometric features.

The outcomes of the current study supply some implications for stakeholders and policymakers in the field of assessment. First, the CALQ conceptual model could be used to portray the underlying components of CBA in qualitative, quantitative, and mixed-method research designs. Second, CALQ could be employed as a valid and reliable instrument to provide educational institutions and policymakers with the practical tool to distinguish EFL teachers according to their CBAL level and facilitate materials preparation to design instructional courses to develop EFL teachers’

CBA knowledge, which could result in the professional development that is influential in students' achievement and performance. Third, CALQ, as a self-assessment tool could be used by teachers to identify their level of CBAL and function as an instrument to investigate their CBA knowledge, skills, and beliefs.

Considering the results of the present study, the subsequent limitations must be taken into consideration. The participants were all volunteers, and it was only feasible for the researchers to choose the respondents according to nonprobability convenience sampling techniques. To develop CALQ, some assessment experts and experienced EFL teachers voluntarily accepted to participate in the study, although there were some criteria for their selection. Their responses were considered the foundation for the questionnaire development. Also due to some limitations, the interviews could be conducted with five experts and 13 EFL teachers. The results could be different being capable of interviewing more participants. Additionally, the researchers were limited to accessing participants to answer the questionnaire who were only from Iran, which would not provide an overall scope of responses. Also, the study only focuses on teachers' role in CBA although the significant role of students in conducting the process of assessment in the classroom context needs to be investigated.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Elham Banisaeed



<http://orcid.org/0000-0002-8460-8093>

Mohammad Hashamdar



<http://orcid.org/0000-0002-1462-8531>

Kobra Tavassoli



<http://orcid.org/0000-0002-8246-8584>

References

- Alyami, A. (2022). *Classroom-based language assessment in a Saudi context: teachers' practices, beliefs and assessment literacy*. [Doctoral dissertation, University College Dublin]. <http://hdl.handle.net/10197/13031>
- Alonzo, D., Leverett, J., & Obsioma, E. (2021). Leading an assessment reform: Ensuring a whole-school approach for decision-making. *Frontiers in Education*, 6(62), 1-11. <https://doi.org/10.3389/educ.2021.631857>
- Aria, D., Sukyadi, D., & Kurniawan, E. (2021). Teacher assessment literacy: Indonesian EFL secondary teachers' self-perceived on classroom-based assessment practice. *English Review: Journal of English Education*, 10(1), pp. 15-26. <https://doi.org/10.25134/erjee.v10i1.5349>
- Bae, J. and Bachman, L. F. (2010). An investigation of four writing traits and two tasks across two languages. *Language Testing* 27 (2), 213-234. <https://doi.org/10.1177/0265532209349470>
- Barnes, N., Fives, H., & Dacey, C. M. (2015). Teachers' beliefs about assessment. https://www.montclair.edu/proflepapes/media/6109/user/Barnes_et_al_2015_teachers_beliefs_about_assessment.pdf
- Best, J. W., & Kahn, J. V. (2006). *Research in education*. Pearson Education.
- Black, P. & William, D. (1998). *Inside the black box: Raising standards through classroom assessment*. Phi Delta Kappan.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81-109. <https://doi.org/10.1017/S0261444803001903>
- Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: what do we need to learn, unlearn, and relearn? *Language Testing in Asia* 10(3), 1-16. <https://doi.org/10.1186/s40468-020-00101-6>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approach* (5th ed.). Sage Publications.
- Crusan, D., Plake, L., & Gebiril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing*, 28, 43-56. <https://doi.org/10.1016/j.asw.2016.03.001>
- Dagdag, J. D., Anoling Jr, O. C., Salviejo, R. P., Pascual, J. F., & Dagdag, J. M. H. (2020). Development of problem-solving efficacy scales in mathematics. *Universal Journal of Educational Research*, 8(6), 2397-2405.

- <http://dx.doi.org/10.13189/ujer.2020.080624>
- Dashti, S. (2019). *EFL teachers' beliefs and practices about classroom assessment: A multiple case study in the context of Kuwait* [Doctoral dissertation, York University]. <https://yorkspace.library.yorku.ca>
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347. <https://doi.org/10.1177/0265532208090156>
- Dehqan, M., & Asadian Sorkhi, S. R. (2020). Pre-service and In-service Teachers' Knowledge and Practice of Assessment Literacy: A Dweller in an Ivory Tower. *Issues in Language Teaching* 9(2), 347-375. <https://doi.org/10.22054/ilt.2021.57203.556>
- Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. Lawrence Erlbaum Associates.
- Farhady, H. (2019). A cross-contextual perspective on EFL teachers' assessment knowledge. *EDU7*, 8(10), 1-19. <https://dergipark.org.tr/en/pub/edu7/issue/51553/665542>
- Farikhah, N., Setyaningsih, E., & Sumardi. (2022). A novice teacher's familiarity with assessment for learning in EFL classroom. *Voice of English Language Society* 6(2). 500-511. <http://dx.doi.org/10.29408/veles.v6i2.6589>
- Field, A. (2018). *Discovering statistics using IBM SPSS, statistics for statistics*. (5th ed.). Sage Publications.
- Fitriyah, I., Massitoh, F., & Widiati, U. (2022). Classroom-based language assessment literacy and professional development need between novice and experienced EFL teachers. *Indonesian Journal of Applied Linguistics*, 12(1), 124-134. <https://doi.org/10.17509/ijal.v12i1.46539>
- Fryer, L. K., Larson-Hall, J., & Stewart, J. (2018). Quantitative methodology. In *The Palgrave handbook of applied linguistics research methodology* (pp. 55-77). Palgrave Macmillan.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
- Fulcher, G., & Davidson, F. (2012). *The Routledge handbook of language testing*. Routledge. <https://doi.org/10.4324/9780203181287>
- George, D., & Mallery, P. (2020). *IBM SPSS statistics 26 step by step: A simple guide and reference*. Routledge.
- Gotch, C. M., & French, B. F. (2014). A systematic review of assessment literacy

- measures. *Educational Measurement: Issues and Practice* 33(2), 14–18. <https://doi.org/10.1111/emip.12030>
- Harrison, V., Kemp, R., Brace, N., & Snelgar, R. (2021). *SPSS for psychologists*. Bloomsbury Publishing. <https://www.bloomsbury.com/uk/spss-for-psychologists-9781352009941>
- Hill, K. & McNamara, T. (2012). Developing a comprehensive research framework for classroom-based assessment – an empirical study. *Language Testing*, 29(3), 395-420. <https://doi/pdf/10.1177/0265532211428317>
- Khany, R. , Azimi Amoli, F. (2016)._Validating an English Language Teacher Professional Development Scale in Iranian EFL Context. *Issues in Language Teaching* 5 (1). 107-134. <https://doi.org/10.22054/ilt.2016.7728>
- Khodashenas, M. R., Khodabakhshzadeh, H., Baghaei, P., & Motallebzadeh, K. (2022). EFL teachers assessment literacy needs inventory: A case of Fulcher’s assessment literacy framework. *Issues in Language Teaching (ILT)* 11(1), 131-156. <https://doi.org/10.22054/ilt.2022.63731.639>
- Kingston, N., Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice* 30(4), 28-37. <https://doi.org/10.1111/j.1745-3992.2011.00220.x>
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative Research: A Guide to Design and Implementation* (4th ed). A Wiley Brand.
- Muñoz, A. P., Palacio, M., & Escobar, L. (2012). Teachers’ beliefs about assessment in an EFL context in Colombia. *Profile Issues in Teachers Professional Development*, 14(1), 143-158. https://www.researchgate.net/publication/317510086_Teachers'_Beliefs_About_Assessment_in_an_EFL_Context_in_Colombia
- Narathakoon, A., Sapsirin, S., & Subphadoongchone, P. (2020). Beliefs and classroom assessment practices of English teachers in primary schools in Thailand. *International Journal of Instruction*, 13(3), 137-156. <https://doi.org/10.29333/iji.2020.13310>
- Nikmard, F., Mohamadi Zenouzagh, Z. (2020). Designing and validating a potential assessment inventory for assessing ELTs’ assessment literacy. *Language Testing in Asia*, 10(8), 1-19. <https://languagetestingasia.springeropen.com/articles/10.1186/s40468-020-00106-1>
- Pallant, J. (2016). *SPSS survival manual* (6th ed.). Allen & Unwin.

- Pill, J., Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30 (3), 381-402. <https://doi.org/10.1177/0265532213480337>
- Rahimi, S. A., Razmjoo, S. A., Sahragard, R., & Ahmadi, A. R. (2021). Development and validation of language assessment literacy scale for high school EFL teachers: A mixed methods approach. *Teaching English Language*, 15(2), 229-126. <https://doi.org/10.22132/TEL.2021.142934>
- Rastegar, B., Zarei, A.A., & Esfandiari, R. (2022). Towards the Development of an Assessment Literacy Questionnaire: The Case of Iranian EFL Teachers. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, 10(2), 9-34. <http://dx.doi.org/10.22049/JALDA.2022.27579.1378>
- Revelle, W. (2020). *Psych: Procedures for personality and psychological research*. Northwestern University, Evanston, Illinois, USA, <https://personality-project.org/r/psych/>
- Riazi, A. M. (2016). *The Routledge Encyclopedia of research methods in applied linguistics*. Routledge. <https://www.routledge.com/The-Routledge-Encyclopedia-of-Research-Methods-in-Applied-Linguistics/Riazi/p/book/9781032098074>
- Sasmaz-Oren, F., & Ormanci, U. (2011). Teacher candidate levels of familiarity with the methods, techniques and tools composing the alternative assessment approaches. *Procedia Social and Behavioral Sciences* 15 (2011) 3476–3483. <https://www.10.1016/j.sbspro.2011.04.321>
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kaplan*, 72(7), 534–39. <https://www.jstor.org/stable/i20404445>
- Tabachnick, B.G. & Fidell, L.S. (2014). *Using multivariate statistics*. (6th ed.). Pearson Inc. <https://ebook.upgrisba.ac.id/ebook/komputer-informasi-referensi-umum/6th-edition-using-multivariate-statistics-pearson/download>
- Tajeddin, Z., Saeedi, Z., & Panahzadeh, V. (2022). English language teachers' perceived classroom assessment knowledge and practice: Developing and validating a scale. *Profile: Issues in Teachers' Professional Development*, 24(2), 247–264. <https://doi.org/10.15446/profile.v24n2.90518>
- Taylor, L. (2013). Communicating the theory, practice, and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403–

412. <https://doi.org/10.1177/0265532213480338>
- Tofighi, S., & Ahmadi Safa, M. (2023). Fairness in Classroom Language Assessment from EFL Teachers' Perspective. *Teaching English as a Second Language Quarterly* (Formerly Journal of Teaching Language Skills), 42(2), 81-110. <https://doi.org/10.22099/tesl.2023.46825.3173>
- Toth, E., Csapo, B. (2022). Teachers' beliefs about assessment and accountability. *Educational Assessment, Evaluation and Accountability* 2022(34), 459-481. <https://doi.org/10.1007/s11092-022-09396-w>
- Unal, A. & Unal, Z. (2019). An examination of K-12 teachers' assessment beliefs and practices in relation to years of teaching experience. *Georgia Educational Researcher* 16(1), 3-23. <https://digitalcommons.georgiasouthern.edu/gerjournal/vol16/iss1/2>
- Watkins, M. W. (2021). *A step-by-step guide to exploratory factor analysis with SPSS*. Routledge. <https://dokumen.pub/a-step-by-step-guide-to-exploratory-factor-analysis-with-spss-1nbsped-0367711117-9780367711115-9780367710316-9781003149347>
- Yamtim, V. & Wongwanich, S. (2014). A study of classroom assessment literacy of primary school teachers. *Procedia - Social and Behavioral Sciences* 116, 2998 – 3004. <https://doi.org/10.1016/j.sbspro.2014.01.696>
- Yan, Z., & Pastore, S. (2022). Are teachers literate in formative assessment? The development and validation of the Teacher Formative Assessment Literacy Scale. *Studies in Educational Evaluation*. <https://doi.org/10.1016/j.stueduc.2022.101183>
- Zhu, X., Raquel, M., & Aryadoust, V. (2019). Structural equation modeling to predict performance in English proficiency tests. In Aryadoust, V., Raquel, M. (Ed.), *Quantitative Data Analysis for Language Assessment Volume II* (pp. 101-126). Routledge. <https://doi.org/10.4324/9781315187808>

Appendices

Appendix A

Interview Questions

1. In your opinion, what kinds of teachers are considered assessment illiterate?
2. What is Validity/ Reliability/ Formative Assessment/ Summative Assessment?
3. What are the different steps in constructing language tests?
4. What is the importance of assessment in education?
5. What is the relationship between assessment and society?
6. Since when the assessment was considered vital?
7. Do you believe in the relationship between experience, education, and assessment literacy?
8. What is the purpose of assessment?

Appendix B

Classroom-Based Assessment Literacy Questionnaire (CALQ)

Gender: Male Female

Years of Experience: Less than 5 years 5 to 10 years More than 10 years

Educational Degree: Diploma Associate degree BA MA

Major: English Non-English

Items	5 Strongly Agree	4 Agree	3 Neither Agree nor Disagree	2 Disagree	1 Strongly Disagree
1 Assessment illiterate teachers are those who are <u>not</u> educated in the English major.					
2 Assessment illiterate teachers are those who are <u>not</u> experienced in teaching and testing English.					
3 Assessment illiterate teachers are those who do <u>not</u> know concepts and methods of language assessment (including knowledge of assessment types, assessment					

theories, assessment rubrics, scoring system, and cut- off score).					
4 Assessment illiterate teachers are those who do <u>not</u> know the importance of contextual factors, individual differences, ethics, commitment, and fairness in assessment.					
5 Assessment illiterate teachers are those who are <u>not</u> able to put their assessment knowledge into practice and construct a test and develop test items.					
6 Assessment illiterate teachers are those who are <u>not</u> able to report test results to students, give feedback, make sound decisions and decide upon remedial courses.					
8 Assessment illiterate teachers are those who do <u>not</u> employ multiple types of assessment and alternative assessment.					
9 Assessment illiterate teachers are those who skip need analysis in assessment.					
10 Assessment illiterate teachers are those who do <u>not</u> align assessment with learning goals.					
11 I am familiar with <u>Assessment</u> .					

12 I am familiar with <u>Alternative Assessment</u> .					
13 I am familiar with <u>Authentic Assessment</u> .					
14 I am familiar with <u>Criterion-Referenced Assessment</u> vs. <u>Norm-Referenced Assessment</u> .					
15 I am familiar with <u>Cut-off Score</u> .					
16 I am familiar with <u>Formative Assessment</u> vs. <u>Summative Assessment</u> .					
17 I am familiar with <u>Qualitative Measurement</u> vs. <u>Quantitative Measurement</u> .					
18 I am familiar with <u>Test Rubrics</u> .					
19 I am familiar with <u>Standardized Testing</u> .					
20 I am familiar with <u>Validation</u> .					
21 <u>Assessment</u> is the systematic process of documenting and using theoretical data to measure knowledge, skills, attitudes and beliefs.					
22 <u>Alternative Assessment</u> refers to procedures and techniques which can be used within the context of instruction and can be easily incorporated into the daily activities of the educational setting.					
23 <u>Authentic Assessment</u> is a form of assessment in which					

students are asked to perform pedagogical tasks that demonstrate meaningful application of essential knowledge and skills.					
24 <u>Criterion- Referenced Assessment</u> measures a student's performance based on mastery of a specific set of skills, whereas <u>Norm-Referenced Assessment</u> measures a student's performance in comparison to the performance of students on the same assessment.					
26 <u>Formative Assessment</u> is conducted at the end of course or unit while <u>Summative Assessment</u> is undertaken during the learning process.					
27 <u>Qualitative Measurement</u> looks for patterns in non-numerical data but <u>Quantitative Measurement</u> involves running statistical analysis on data that has numerical values.					
29 <u>Standardized Testing</u> are explicit sets of criteria used for assessing a particular type of work or performance and provides more details than a single grade or mark.					
30 <u>Validation</u> in assessment means the quality review of the assessment tools, processes, practices and judgments.					

31 I am familiar with the steps of test construction.					
32 Determining the function and the form of the test is among major steps of test construction.					
33 Determining the content of the test is among major steps of test construction.					
36 Preparing the items is among major steps of test construction.					
37 Reviewing the items is among major steps of test construction.					
41 Assessment is the most important part of education.					
42 Without assessment, education can achieve the predetermined goals.					
43 The partial philosophy of assessment is improvement and development.					
45 The philosophy of assessment is providing feedback and modifying problems in learning and teaching.					
46 Assessment plays the role of a criterion in society to select more competent applicants in different fields.					
47 There are no alternatives for assessment.					
48 The idea of assessment cannot be omitted, but traditional assessment might					

be replaced by dynamic assessment.					
49 Assessment has always been vital.					
50 Assessment has recently been considered crucial without the advent of modern types of assessment.					

Appendix C Tables and Figures

Table 1

Standardized Scores for Items (Checking Univariate Outliers)

Items	Min	Max	Items	Min	Max	Items	Min	Max
S3	-2.13	1.55	S31	-1.53	2.11	S38	-1.43	1.40
S4	-2.09	1.50	S21	-1.41	2.02	S39	-1.52	1.35
S5	-2.11	1.51	S22	-1.43	2.00	S40	-1.48	1.38
S7	-1.37	1.42	S23	-1.50	2.08	S6	-1.51	2.03
S8	-2.13	1.56	S24	-1.50	1.99	S41	-1.52	2.05
S9	-2.03	1.49	S25	-1.39	1.37	S42	-1.45	2.04
S10	-2.08	1.47	S26	-1.49	2.05	S44	-1.50	1.40
S11	-1.49	2.07	S27	-1.50	2.08	S46	-1.42	2.02
S12	-1.51	2.07	S28	-1.39	1.36	S49	-1.44	2.04
S13	-1.47	2.05	S29	-1.44	2.01	S50	-1.49	2.04
S14	-1.47	2.03	S30	-1.53	2.14	S1	-1.57	2.08
S15	-1.47	2.05	S32	-1.54	2.08	S2	-1.52	2.11
S16	-1.46	2.07	S33	-1.52	2.05	S43	-1.50	2.06
S17	-1.57	2.14	S34	-1.46	1.51	S45	-1.45	2.00
S18	-1.49	2.01	S35	-1.51	1.41	S47	-1.48	2.03
S19	-1.46	2.08	S36	-1.50	2.07	S48	-1.52	2.13
S20	-1.46	1.99	S37	-1.52	2.06			

Table 2

Descriptive Statistics for Mahalanobis Distances (Checking Multivariate Outliers)

	N	Minimum	Maximum
Mahalanobis Distance	318	29.14	72.20
Critical Value of Chi-square (.001, 50)	88.66		

Table 3

Skewness and Kurtosis, and Mardia Indices of Univariate and Multivariate Normality

Item	Skewness	Kurtosis	Item	Skewness	Kurtosis	Item	Skewness	Kurtosis
S48	0.241	-0.489	S5	-0.169	-0.715	S26	0.225	-0.565
S47	0.177	-0.734	S8	-0.175	-0.598	S27	0.126	-0.702
S45	0.268	-0.698	S9	-0.191	-0.648	S31	0.124	-0.775
S43	0.127	-0.753	S10	-0.130	-0.816	S20	0.156	-0.907
S2	0.086	-0.810	S37	0.168	-0.732	S19	0.109	-0.765
S1	0.137	-0.622	S36	0.188	-0.668	S18	0.302	-0.632
S6	0.166	-0.716	S33	0.203	-0.680	S17	0.176	-0.568
S41	0.112	-0.747	S32	0.093	-0.772	S16	0.289	-0.601
S42	0.223	-0.626	S30	0.213	-0.601	S15	0.231	-0.709
S46	0.187	-0.789	S29	0.277	-0.650	S14	0.184	-0.736
S49	0.283	-0.692	S21	0.324	-0.659	S13	0.195	-0.702
S50	0.187	-0.747	S22	0.200	-0.876	S12	0.167	-0.657
S3	-0.175	-0.592	S23	0.262	-0.640	S11	0.206	-0.688
S4	-0.266	-0.631	S24	0.227	-0.735	Mardia		-1.19

Table 4

Component Correlation Matrix (Second EFA)

Component	1	2	3	4	5	6
1	---					
2	.306	---				
3	-.274	-.303	---			
4	-.290	-.306	.276	---		
5	-.318	-.312	.293	.302	---	
6	-.298	-.318	.336	.319	.316	---

Table 5

Watkins' Parallel Analysis

Factors	Eigenvalues			Factors	Eigenvalues		
	Simulated	Initial	Decision		Simulated	Initial	Decision
1	1.759	10.946	Keep	22	0.934	0.480	Drop
2	1.673	3.773	Keep	23	0.906	0.472	Drop
3	1.607	2.755	Keep	24	0.880	0.454	Drop
4	1.550	2.544	Keep	25	0.857	0.439	Drop
5	1.499	2.432	Keep	26	0.831	0.435	Drop
6	1.455	2.258	Keep	27	0.806	0.422	Drop
7	1.415	0.701	Drop	28	0.782	0.411	Drop
8	1.371	0.685	Drop	29	0.758	0.394	Drop
9	1.333	0.666	Drop	30	0.734	0.389	Drop
10	1.295	0.645	Drop	31	0.709	0.373	Drop
11	1.260	0.626	Drop	32	0.685	0.369	Drop
12	1.227	0.618	Drop	33	0.661	0.356	Drop
13	1.194	0.617	Drop	34	0.637	0.337	Drop
14	1.163	0.584	Drop	35	0.612	0.334	Drop
15	1.131	0.571	Drop	36	0.586	0.329	Drop
16	1.099	0.562	Drop	37	0.562	0.322	Drop
17	1.069	0.543	Drop	38	0.536	0.302	Drop
18	1.042	0.537	Drop	39	0.507	0.292	Drop
19	1.015	0.518	Drop	40	0.475	0.269	Drop
20	0.987	0.499	Drop	41	0.438	0.246	Drop
21	0.960	0.494	Drop				

Table 6*KMO and Bartlett's Test (Second EFA)*

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.937
Approx. Chi-Square		6275.993
Bartlett's Test of Sphericity	Df	820
	Sig.	.000

Table 7*Total Variance Explained (Second EFA)*

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	10.94	26.697	26.697	10.483	25.567	25.567	6.041	14.733	14.733
2	3.773	9.204	35.900	3.318	8.092	33.660	3.422	8.345	23.078
3	2.755	6.719	42.619	2.313	5.641	39.301	3.171	7.735	30.814
4	2.544	6.205	48.824	2.071	5.051	44.352	3.168	7.727	38.541
5	2.432	5.932	54.756	1.962	4.785	49.136	3.165	7.719	46.260
6	2.258	5.506	60.262	1.760	4.292	53.429	2.939	7.169	53.429
7	.701	1.709	61.971						
8	.685	1.672	63.643						
9	.666	1.624	65.267						
10	.645	1.574	66.841						
11	.626	1.526	68.366						
12	.618	1.508	69.875						
13	.617	1.505	71.379						
14	.584	1.425	72.804						
15	.571	1.392	74.196						
16	.562	1.371	75.568						
17	.543	1.325	76.892						
18	.537	1.309	78.201						
19	.518	1.263	79.464						
20	.499	1.218	80.682						
21	.494	1.204	81.886						
22	.480	1.171	83.057						
23	.472	1.151	84.209						
24	.454	1.107	85.316						
25	.439	1.072	86.387						
26	.435	1.062	87.449						
27	.422	1.029	88.478						
28	.411	1.003	89.480						
29	.394	.962	90.442						
30	.389	.948	91.390						
31	.373	.910	92.300						
32	.369	.901	93.201						
33	.356	.867	94.068						
34	.337	.823	94.891						
35	.334	.815	95.706						
36	.329	.803	96.509						
37	.322	.786	97.295						
38	.302	.736	98.032						



39	.292	.712	98.743
40	.269	.656	99.399
41	.246	.601	100.000

Extraction Method: Principal Axis Factoring.

Figure 1
Six Measurement Models of CALQ

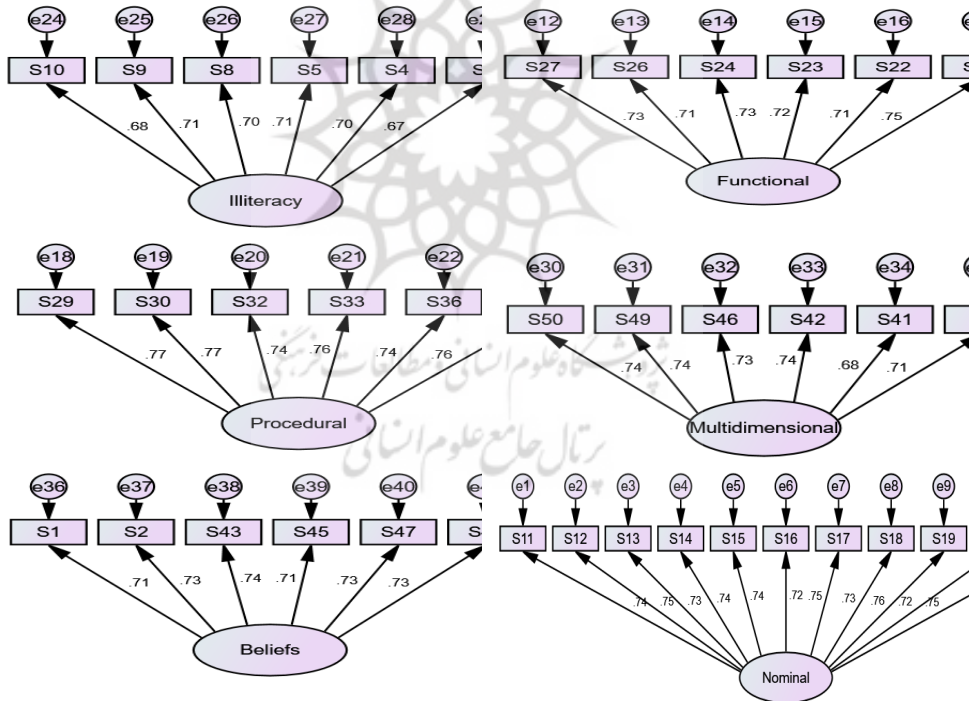


Table 8
Fit Indices for six Measurement Models

	Indices	Index	Criteria	Fit
Absolute	X ²	450.78	---	---
	df	773	---	---
	p	1.00	> .05	Good Fit
	X ² Ratio	.583	< 3	Good Fit
	SRMR	.024	<.10	Good Fit
	RMSEA	.000	<.05	Good Fit
	CI	[.000,.000]	<=.10	Good Fit
	PCLOSE	1.00	>.05	Good Fit
	GFI	.938	>=.90	Good Fit
Incremental	RFI	.927	>=.90	Good Fit
	TLI	1	>=.90	Good Fit
	CFI	1	>=.90	Good Fit
	NFI	.931	>=.90	Good Fit
	IFI	1	>=.90	Good Fit
Hoelter (Sampling Adequacy)		590	>200	Adequate

