

Improving Language Assessment Literacy for In-Service Saudi EFL Teachers

Faten Ali Salami

King Abdulaziz University, Jeddah, Saudi Arabia
Corresponding Author: falisalami@stu.kau.edu.sa

Rasha Mastour Alharthi

King Abdulaziz University, Jeddah, Saudi Arabia

Received:0/4/2022

Accepted:09/01/2022

Published:09/24/2022

Abstract

A teacher's ability to create and implement high-quality assessments is a key component of assessment literacy. Therefore, improving teachers' language assessment literacy (LAL) is one of the most important factors in assuring the quality of language testing and assessment (LTA). Thus, this study aimed to investigate the perceived level of LAL among Saudi in-service English as a foreign language (EFL) teachers. Additionally, it aimed to identify the aspects of LAL that in-service EFL teachers wish to improve through further professional assessment training. Two research questions were formulated to achieve the study aims; a) what is the perceived level of LAL among Saudi in-service EFL teachers? b) What aspects of LAL do in-service EFL teachers wish to improve through further professional training? A questionnaire was administered to collect the data. Fifty EFL female instructors at King Abdulaziz University responded to the online questionnaire. To analyze the data, SPSS software was used. Regarding teachers' perceived level of LAL, the results indicated that Saudi EFL teachers lack a well-developed knowledge of assessment literacy. The findings showed that the majority of the teachers received either no training or basic training in LAL. In regard to teachers' training needs, the findings highlighted that there is a need for training in nearly all the aspects of LAL, with priority being given to different content areas of LTA. Finally, the study recommended providing Saudi EFL teachers with sufficient professional language assessment training.

Keywords: Content of assessment, language assessment literacy, language testing, Saudi EFL teachers.

Cite as: Salami, F. A. , & Alharthi, R. M. (2022). Improving Language Assessment Literacy for In-Service Saudi EFL Teachers *Arab World English Journal*, 13 (3) 536 -554.

DOI: <https://dx.doi.org/10.24093/awej/vol13no3.35>

Introduction

It is widely accepted that one of the most challenging aspects of a teacher's job is assessing students' performance (Huang & He, 2016). Assessment is an integral part of the teaching and learning process. Language learning assessment is not only a means to evaluate students' progress and improve their achievement but also a way to enhance the learning process (Earl, 2003). Since teachers play a central role in assessment, it is essential that they have a sufficient level of language assessment literacy (LAL) so that they may effectively and fairly assess their students' learning. (DeLuca & Klinger, 2010; Koh et al., 2018; Koh, 2011).

Stiggins (1991) defined assessment literacy as teachers' knowledge of both the underlying principles and the skills needed to evaluate student learning. Furthermore, those who are assessment literate should know what to assess, how to assess, what challenges to expect, and how to overcome them (Stiggins, 1995). This means assessment-literate teachers should be knowledgeable about how to use different kinds of assessment to assist students, how to effectively and accurately convey assessment results, how to give feedback to their students, and how to use the results to improve the learning process.

Due to its significance, an increasing number of studies have been conducted that attempt to measure teachers' assessment literacy in various contexts and aspects (Berry et al., 2019; Fulcher, 2012; Hasselgreen et al., 2004; Lam, 2015; Mendoza, 2009; Vogt & Tsagari, 2014). However, few studies have explored the language assessment literacy of EFL teachers in Saudi Arabia (Hakim, 2015; Umer et al., 2018). Therefore, the current study aimed to investigate the perceived level of LAL among Saudi in-service EFL teachers. Additionally, it aimed to identify the aspects of LAL that in-service EFL teachers wish to improve through further professional assessment training.

Literature Review

Language Assessment Literacy

In a broader sense, the term "language assessment literacy" refers to educators' and stakeholders' ability to understand what, why and how to assess, the possible problems with assessment, and the negative consequences of inaccurate assessment (Stiggins, 2002). Fulcher's (2012) definition of LAL is based on his research of language teachers' assessment training needs. Fulcher (2012) concluded that LAL refers to:

The knowledge, skills, and abilities required to design, develop, maintain, or evaluate large-scale standardized and/or classroom-based tests, familiarity with test processes, and awareness of principles and concepts that guide and underpin practice, including ethics and codes of practice (Fulcher, 2012, p. 125).

Similarly, Davies (2008) noted that LAL has three components: skills, knowledge, and principles. *Skills* refers to the practical application of knowledge in test analysis and design, and *knowledge* refers to the "relevant background in measurement and language description" (p. 328). *Principles* encompasses social-related context issues, such as test fairness and test ethics.

Conversely, Newfields (2006) has a different perspective on the concept of assessment literacy. He defined LAL according to the viewpoints of stakeholders, such as students, teachers,

and test developers. For students, LAL refers to knowing how to perform well on exams. For teachers, it is associated with the ability to grade students accurately and ethically. For test developers, every aspect of their work is related to assessment literacy. Similarly, Pill and Harding (2013) viewed LAL as a “repertoire of competences that enable an individual to understand, evaluate and, in some cases, create language tests and analyze test data” (p. 282). They noted that different kinds of expertise or specializations require different levels of assessment literacy. For instance, the level of assessment literacy required by practitioners, such as test designers, teachers, and school principals, is quite different from the level required for non-practitioners, such as test takers.

Teachers’ Language Assessment Literacy

According to Harding and Kremmel (2016), because language teachers are the key users of language assessment, they must be familiar with and proficient in its practices. EFL teachers play an essential role in preparing and using language assessments, and a great deal of research has examined language teachers’ perceptions, needs, and knowledge of assessment. On the one hand, some studies have found that teachers lack knowledge related to language assessment (Fulcher, 2012; Lam, 2015). On the other hand, some studies claim teachers may be aware of the skills, knowledge, and concepts required in assessment, but they do not put them into practice (Shim, 2009; Muoz et al., 2012). A study conducted by Shim (2009) used a questionnaire and interviews with some of the study’s participants to evaluate instructors’ beliefs and practices surrounding classroom-based English language evaluation. The findings showed that teachers were assessment literate and aware of the principles of assessment and testing; however, they did not put all the principles into practice.

Similarly, Muoz et al. (2012) investigated sixty-two Colombian teachers’ opinions of assessment. They found a gap between teachers’ attitudes and the practices that needed to be addressed through teacher training courses. Using teacher survey responses and focus group interviews, Qassim (2008) examined the factors that influenced teachers’ assessment procedures in public secondary schools in Qatar. The majority of instructors stated that although they were assessment literate, they could not connect their knowledge to their practice. The study presented the following reasons for this disconnect: curriculum workload, teaching time, and class size, all of which have been identified as variables that hamper the use of various forms of assessment in the classroom. These factors should be considered when training teachers so they can more effectively apply what they have learned about LAL.

EFL Teachers’ Language Assessment Training Needs

Since many studies have confirmed the importance of formal assessment training for EFL teachers, knowing where and how to start a process of professional assessment training is crucial for ensuring a positive outcome. Fulcher (2012) used an online survey to identify the training needs of foreign language teachers. His findings implied that language teachers are aware of the aspects of language assessment that are not currently addressed in available training materials. Therefore, accessible and practical materials are needed to help them improve their assessment literacy. Mendoza (2009) found that teachers frequently and inappropriately use summative rather than formative assessments. They also lack knowledge of critical aspects of language assessment, such as the different types of language assessment, providing effective feedback to students, and ethical

issues related to testing and assessment. Similarly, DeLuca and Klinger (2010) discovered that while 288 Canadian EFL teachers were aware of summative evaluations, they were unfamiliar with formative assessment. Therefore, DeLuca and Klinger emphasized the need for direct education in the development of assessment literacy in teachers.

Furthermore, it is important that the type of training and professional development offered is tailored to teachers' needs. Hasselgreen et al. (2004) surveyed teachers, teacher trainers, and experts in language assessment from thirty-seven European countries. They were asked to state the extent to which they had received formal training and the degree to which they believed there was a need for training. The survey addressed four aspects of assessment: classroom-focused assessment, purposes of assessment, content and concepts of assessment, and external testing and exams. The results revealed that for classroom-focused assessment, teachers often had little training in using ready-made assessments, delivering comments, or employing informal or continuous assessment. Despite their perception that there was no need for training in using ready-made exams, they did identify a need for training in test preparation, evaluating findings, offering feedback, and using self-assessment and portfolios. Regarding the areas of selecting the purposes of assessment, granting final certificates, and selecting the students' needs, teachers reported that they frequently performed these aspects without having received any training.

A mixed methods study by Vogt and Tsagari (2014) explored the same aspects of language assessment that were investigated by Hasselgreen et al. (2004). However, in order to obtain accurate results regarding teachers' assessment training needs, Vogt and Tsagari focused only on regular teachers. To some extent, the training needs of EFL teachers in the various areas of LTA identified in this study confirm Hasselgreen et al.'s findings. The results showed a general need for training in both surveys, and teachers, in particular, reported a need for basic or advanced training in almost all topics covered in the survey. For example, they reported that they received some training in 'giving feedback', preparing self/peer-assessments, and in various content areas, however, they asked for more training in these aspects. The results also confirmed that teachers had no training in some areas such as 'preparing classroom tests' and 'giving grads'. Thus, language assessment training for EFL instructors has been confirmed (Vogt & Tsagari, 2014).

Studies on Language Assessment Literacy in the Saudi Context

Limited research has been conducted in the Saudi EFL context on the different aspects of LAL (Hakim, 2015; Umer et al., 2018). Hakim (2015) conducted a quantitative study to investigate teachers' awareness levels of various assessment tools and their perceptions of them. The results indicated that teachers had a clear knowledge of assessment tools, as they had participated in in-house and in-service workshops and training programs. However, it was discovered that novice teachers did not have a sufficient understanding of practical approaches to assessment, unlike experienced instructors who had a greater comprehension of LAL and successfully modified their assessment practices. It is reasonable to conclude that new language instructors, in particular, should be encouraged to improve their assessment literacy through training courses aimed at providing them with a clear and thorough understanding of sound assessment components, concepts, ideologies, and practices.

Additionally, Umer et al. (2018) explored how Saudi EFL teachers' assessment practices aligned with recommended practices. The results revealed that most assessment tasks focused on lower order learning and that EFL teachers lacked knowledge of grading and preparing a comprehensive answer key. This study confirmed the gap between assessment theory (recommended, sound assessment principles) and teachers' actual assessment practices, highlighting the need for professional training in assessment literacy.

Previous studies conducted in Saudi Arabia confirmed the importance of language assessment training for EFL instructors. However, to the best of our knowledge, no study has been conducted to identify which aspects of LAL need to be improved based on teacher feedback. Therefore, this study intends to bridge the gap in the existing literature by answering the following research questions to gain a deeper understanding of LAL among Saudi EFL teachers:

1. What is the perceived level of LAL among Saudi in-service EFL teachers?
2. What aspects of LAL do in-service EFL teachers wish to improve through further professional training?

Methodology

Research Design and Instruments

In the field of education, descriptive research is used to investigate the *what* aspect of any phenomenon (Cohen et al., 2007). Therefore, this study followed a descriptive study design and used a quantitative questionnaire that was administered to female EFL teachers in Saudi Arabia. The questionnaire was adopted from Vogt and Tzagari (2014), which was originated by Hasselgreen et al. (2004). The questionnaire consists of three parts (see Appendix A). The first part includes respondents' demographic data, such as age and level of education. The second part contains thirty items, thematically grouped as follows: (a) Language Testing and Assessment (LTA), (b) purpose of testing, and (c) content and concept of LTA. Each theme is divided into two sections: one for perceived assessment training and the other for assessment training needs. The third part includes four open-ended questions that aim to provide deeper insight into teachers' assessment competencies and training needs. The three-point Likert scale in Vogt and Tzagari's study (no/a little/more advanced) was used for the closed-ended questions and was slightly modified to (no/yes, basic training/ yes, advanced training). This modification was made to overcome the shortcoming of the original questionnaire.

Participants

The population of this study was female Saudi EFL teachers who work at the English Language Institute (ELI) at King Abdulaziz University in Jeddah, Saudi Arabia. The sample was fifty female EFL teachers from the ELI who participated in this study. Their ages ranged from 25 to 50 years old. They were of different educational levels: BA (38.8%), MA (46.9%), and doctoral (14.3%). The participants were randomly chosen to volunteer for this study. Random sampling was used to minimize any subjective factors that could influence the study (Dörnyei, 2007).

Procedure and Method of Analysis

The data collection process took place from November 15 to November 30, 2021. The questionnaire was designed using Google Forms. Before administering the questionnaire, a pilot

study was conducted with four teachers to ensure the clarity of the items. The teachers reported that the wording of the instructions for the three subscales was unclear, so amendments were made to address this confusion. The questionnaire was sent to ELI faculty members via email and WhatsApp. After collection, the data were organized using Excel and then passed through SPSS software for analysis. Descriptive statistics were run to explain the frequencies of the responses in the questionnaire, the mean, and the standard deviation.

Ethical Considerations

Ethical issues in data collection were taken into consideration. Approval was requested and granted by the ELI at King Abdulaziz University (see Appendix B). All teachers participated voluntarily. They were assured that their identities and responses would remain anonymous and that their data would only be used for research purposes.

Results

Demographic Data

In Table 1, frequency statistics show that the majority of participants (74%) had received training in language assessment. Moreover, among the fifty EFL instructors surveyed, the majority (24%) have master's degrees, 19% have bachelor's degrees, and 7% have doctoral degrees. The age of the participants ranged from 25 to 50. Responses with the highest percentage were from participants aged 30 to 40, whereas the responses with the lowest percentage were from those aged 40 to 50 (see Table 1).

Table 1. *Participants' Demographic Information*

	Demographic information	N	Percentage
Receiving training	Yes	37	74%
	No	13	26%
Age	25–30	22	44%
	30–40	23	46%
	40–50	5	10%
Level of education	Bachelor's	19	38%
	Master's	24	48%
	Doctoral	7	14%

Closed Response Items

The multi-item scales in the questionnaire were tested using Cronbach's alpha to ensure their reliability (see Table 2). According to Dörnyei (2007), Cronbach's alpha must be above .70. The reliability of the three scales was excellent: language testing and assessment training ($\alpha = .86$), training in purpose of language assessment ($\alpha = .80$), and training in concept and content of language assessment ($\alpha = .84$). Total reliability for the entire questionnaire was $\alpha = .91$ (see Table 2).

Table 2. *Reliability Analysis*

Scales	No. of items	Cronbach's alpha
Language testing and assessment	14	.86
Training in purpose of language assessment	8	.80
Training in concept and content of language assessment	8	.84

Comparing the means of the three scales of perceived training in the LAL section, the highest mean score was for the content and concept of language assessment scale $M = 1.87$; Table 3). This indicates that Saudi EFL teachers have knowledge of the items on the aforementioned scale. The lowest mean score was for the LTA scale $M = 1.73$, which means that Saudi EFL teachers have little knowledge of this area of testing and assessment (see Table 3).

Table 3. *Saudi EFL Teachers' Perceived Training in LAL*

Scales	Mean
Language testing and assessment scale	1.73
Purpose of testing scale	1.74
Content and concepts of language assessment scale	1.87

For the training needs section, the highest mean was for the LTA scale $M = 2.08$, which indicates that Saudi EFL teachers need general training on LTA. The lowest mean score was for the purpose of testing scale $M = 1.96$, which shows that Saudi EFL teachers do not require much training on the purposes of testing (see Table 4).

Table 4. *Saudi EFL Teachers' Training Needs in LAL*

Subscales	Mean
Language testing and assessment scale	2.08
Purpose of testing scale	1.96
Content and concepts of language assessment scale	1.97

Training in Language Testing and Assessment

On the LTA scale, the majority of participants had received training in “giving feedback to students based on information from assessments” $M = 1.98$. However, participants desired additional training on both the basic (14%) and advanced (22%) levels. “Preparing classroom tests” $M = 1.84$ and “using self- or peer assessment” $M = 1.86$ were the second most common areas in which participants had received training. However, additional training is needed in both areas, as the averages were $M = 2.06$ and $M = 2.38$, respectively. Few participants had received training in “giving feedback to students based on information from computer-based tests” $M = 1.42$. Therefore, over two-thirds of participants need basic and advanced training in this aspect of language assessment (see Table 5).

Table 5. Respondents' Perceptions of Training Received and Needed in Language Assessment and Testing in LTA

Scale	Items	Percentage			Mean (SD)	
		Received	Needed		Received	Needed
Training in LTA	Preparing classroom tests	30	34	No	1.84(0.65)	2.06(0.86)
		56	26	Basic		
		14	40	Advanced		
	Using ready-made tests from textbook packages or from other sources	36	40	No	1.78(0.67)	1.90(0.83)
		50	30	Basic		
		14	30	Advanced		
	Giving feedback to students based on information from performance assessments	24	28	No	1.98(0.68)	2.16(0.84)
		54	28	Basic		
		22	44	Advanced		
	Giving feedback to students based on information from computer-based tests	64	30	No	1.42(0.60)	2.10(0.83)
		30	30	Basic		
		6	40	Advanced		
	Using self- or peer assessment	32	12	No	1.86(0.70)	2.38(0.69)
		50	38	Basic		
		18	50	Advanced		
	Using an informal, continuous, non-test type of assessment	38	40	No	1.74(0.66)	1.94(0.86)
		50	26	Basic		
		12	34	Advanced		
	Using the European Language Portfolio, an adaptation of it or some other portfolio	62	32	No	1.52(0.73)	2.06(0.84)
		24	30	Basic		
		14	38	Advanced		

Training in the Purpose of Testing

On the purpose of testing scale, the majority of participants had received training in the areas of “giving grades” $M = 2.04$ and “finding out what needs to be taught/learned” $M = 2.10$. However, the least developed areas seemed to be “placing students into courses, programs, etc.” $M = 1.38$ and “awarding final certificates (from school/program; local, regional, or national level)” $M = 1.46$. For training needs, the participants reported a strong need for training in “placing students into courses, programs, etc.” $M = 2.02$ and “finding out what needs to be taught/ learned” $M = 2.02$ (see Table 6).

Table 6. Respondents' Perceptions of Training Received and Needed in Purposes of Testing

Scale	Items	Percentage			Mean (SD)	
		Received	Needed		Received	Needed
Training in purpose of testing	Giving grades	28	46	No	2.04(0.78)	1.82(0.85)
		40	26	Basic		
		32	28	Advanced		
	Finding out what needs to be taught/ learned	16	38	No	2.10(0.64)	2.02(0.89)
		58	22	Basic		
		26	40	Advanced		
	Placing students into courses, programs, etc.	72	38	No	1.38(0.66)	2.02(0.89)
		18	22	Basic		
		10	40	Advanced		
	Awarding final certificates (from school/program; local, regional or national level)	68	32	No	1.46(0.73)	2.00(0.80)
		18	36	Basic		
		14	32	Advanced		

Training in Concept and Content of Assessment

As mentioned above, participants received the greatest amount of training on the aspects of this scale compared to the previous two scales. Even though teachers seemed to be familiar with assessing traditionally tested linguistic skills, such as speaking, writing, grammar, and vocabulary, they still need training in testing and assessing productive $M = 1.98$ and micro-linguistic skills $M = 1.96$ (see Table 3). Testing and assessing integrated skills were areas in which participants needed the most training $M = 2.02$ (see Table 7).

Table 7. Respondents' Perceptions of Training Received and Needed in Content and Concept of Assessment

Scale	Items	Percentage			Mean (SD)	
		Received	Needed		Received	Needed
Training in concept and content of the assessment	Testing/assessing of receptive skills (reading/listening)	36	42	No	1.86(0.75)	1.92(0.87)
		42	24	Basic		
		22	34	Advanced		
	Testing/assessing of productive skills (speaking/writing)	32	38	No	1.90(0.73)	1.98(0.86)
		46	26	Basic		
		22	36	Advanced		
	Testing/assessing of micro-linguistic aspects (grammar/vocabulary)	28	42	No	1.94(0.71)	1.96(0.90)
		50	20	Basic		
		22	38	Advanced		
	Testing/assessing of Integrated language skills	38	38	No	1.78(0.70)	2.02(0.89)
		46	22	Basic		
		16	40	Advanced		

Open Response Items

At the end of the questionnaire, participants were asked the following four open-ended questions to gain deeper insights into their perceived level of LAL and their language assessment training needs.

Q1: During your pre-service training, what did you learn about language testing and assessment?

In this question, participants were asked about their pre-service assessment training. Approximately half of the participants (54%) stated that they did not receive any training prior to starting their jobs. The remaining 46% said that they had received some pre-service assessment training. Of this number, 22% mentioned that they had learned about different types of tests and assessments: "I learned about the type of assessment and when to use them depending on what we want to measure." Another 12% of participants declared that they had learned about assessing language skills: "I learned the assessment of listening, speaking, reading and writing, and integration of two or more skills." Few participants reported that they had received training on how to construct tests and follow rubrics when marking performance tests.

Q2: Have you received in-service training in language assessment? If yes, what was the focus of this training?

In this question, participants were asked whether they had received in-service assessment training. Approximately half of the participants (52%) admitted that they had not received any assessment training while working as EFL teachers. Thirty percent stated that they had taken part in in-service training on test types, formatively and summatively assessing receptive and productive skills, and purposes of testing. Ten percent indicated that they had been trained on marking students' tests and using analytical scales for speaking and writing tests. One participant mentioned that she had received training on using technology in online assessments.

Q3: Have you received training in the use of technology in language assessment? If yes, what was the focus of this training?

This question examined whether participants had been trained in the use of technology in language assessment. The data showed that 78% of participants had not received training on integrating technology into language assessment. Twenty percent of participants stated that they had been trained in using the Blackboard platform to administer tests, especially during the Covid-19 pandemic: "Yes, how to make tests on Blackboard and other applications." One participant reported that she had been trained on the automated coring method.

Q4: Which aspects of language assessment literacy (LAL) do you need to have training on and why?

In this question, participants were asked to identify the aspects of LAL on which they desired future training. There were five aspects of LAL from which participants could choose: purpose of testing, preparing and administering language tests, scoring and rating, giving feedback to students, and assessment policy and local practices. Participants could choose more than one response and were asked to justify their choices. These aspects of LAL were proposed by Kremmel and Harding (2020). Table 8 below shows that the majority of participants chose to have training on preparing and administering language tests and giving feedback to students (54% and 52%, respectively). Forty-eight percent of participants chose to have training on scoring and rating, while thirty-eight percent wanted further training on assessment policy and local practices. Surprisingly, the least frequently chosen aspect was purpose of testing; only 26% wished to receive training on it.

Table 8. *Language Assessment Training Needs for EFL Teachers*

LAL Aspects	No. of Participants	Reasons
Purpose of testing	13	T1 <i>"To be able to prepare for language tests and be aware of its purpose."</i> T2 <i>"I have little knowledge on the testing purposes."</i> T3 <i>"Because most of the teachers write the test just because it is part of their job and do not pay attention to the purpose of test."</i>
Preparing and administering language tests	27	T1 <i>"I still need to be familiar with how to prepare tests which suit different capabilities."</i> T2 <i>"Because as a teacher, I have not been trained on how to make tests. Even though I work in a place that doesn't involve classroom teachers in constructing the tests, but I need this knowledge for my professional development."</i> T3 <i>"To improve my skills in terms of preparing valid and reliable tests."</i>
Scoring and rating	24	T1 <i>"I don't know how to correct students' essays."</i> T2 <i>"I chose scoring because that differs from campus to another."</i> T3 <i>"To be more specific in scoring and not just deducting grades."</i>
Giving feedback to students on their tests	26	T1 <i>"Because I believe that feedback is the most important aspect to help students move forward toward their goal."</i> T2 <i>"I need to know more about them."</i> T3 <i>"I believe it is what I need."</i>
Assessment policy and local practices	19	T1 <i>"Because teaching at different institutions would confuse the teacher on how to assess his/her students since different institutions have different goals and visions."</i>

Discussion

The Level of LAL Perceived by Saudi In-Service EFL Teachers

In this quantitative study, the findings showed that on the training in LTA scale, teachers were well trained in giving feedback to students based on information gleaned from assessments. This finding aligns with Vogt and Tzagari (2014), who found that most EFL teachers from different European countries had received either basic or advanced training in this area. It can be noted that giving feedback is a critical skill, as many aspects of teaching and learning can be developed based on this feedback. For example, feedback gives teachers, students, and test developers insights into the strengths and weaknesses of the course as a whole and the assessments in particular.

On the same scale, teachers showed a high level of training in "preparing classroom tests." This finding is consistent with that of Hasselgreen et al. (2004), who also found that teachers had adequate knowledge in this area. Additionally, teachers in this study received a high level of training in "using self- or peer assessment." This finding is in contrast to that of Hasselgreen et al., who found that teachers lack knowledge in this same area. This may be due to the differences between samples of the two studies, as Hasselgreen et al. focused on experts and teacher trainers who may not be as familiar with classroom-based tests as regular teachers. An unexpected result in the current study was that teachers lacked knowledge of giving feedback based on information collected from computer-based tests. Technology has become essential in different aspects of

language teaching; its use is not optional. Thus, teachers reported a significant need for training in this area.

On the training for the purpose of testing scale, the results indicated that Saudi EFL teachers were knowledgeable in the areas of “giving grades” and “finding out what needs to be taught/learned.” These results are similar to those of Vogt and Tsagari (2014), who reported that teachers had received no training in “giving grades” and that they desired more training in this aspect. It also supported Umer et al. (2018) claim that Saudi teachers lacked grading knowledge. An interesting result was the selection of “placing students into courses, programs, and awarding final certificates” as the area where EFL teachers felt they had received the least amount of training in language assessment. Since purposes of assessment are primarily the domain of administrators rather than regular teachers, it is understandable that Saudi EFL teachers lack knowledge in this area.

The last scale focused on training in concept and content of the assessments. The findings showed that teachers were familiar with assessing language skills (writing, speaking, and reading) and micro-linguistic skills (grammar and vocabulary). However, they reported a significant need for more basic and advanced training in these areas. These results are in line with previous studies by Hasselgreen et al. (2004) and Lan and Fan (2019), who found that teachers wished to improve their abilities in assessing traditional language skills.

The open-ended questions revealed that Saudi EFL teachers had received pre-service training in “preparing different types of assessments in testing/assessing different language skills.” Similarly, Hakim (2015) found that Saudi EFL teachers were knowledgeable about different types of assessments. Moreover, teachers participating in in-service training received more in-depth instruction about specific assessment types, such as formative and summative. The most noteworthy result was that teachers lacked significant knowledge of “using technology in online assessment.” For example, few teachers reported receiving training on using the Blackboard platform to administer tests to their students. Covid-19 has transformed education from a traditional to an online learning environment; therefore, the use of online assessments has become a necessity that every teacher should be trained on.

The Aspects of LAL That Teachers Wished to Improve Through Further Professional Training

Teachers in the study reported a stronger need for future training on aspects of the LTA subscale compared to the other two subscales. For example, teachers perceived “using self- or peer assessment” as their greatest training need. This result corresponds with that of Vogt and Tsagari (2014), who also found that most teachers desire training in using self- or peer assessment. This can be attributed to the fact that participants viewed this type of assessment as a new field of LTA. Teachers identified “preparing classroom tests” as the second area in which they need further training. This finding supports that of Umer et al. (2018), who found a lack of test preparation knowledge among Saudi EFL teachers; thus, they recommended further training in this area. On the same subscale, teachers did not wish to receive training in “using ready-made tests,” which aligns with the findings of both Hasselgreen et al. (2004) and Vogt and Tsagari. Their studies

showed that “using ready-made tests” was teachers’ least desired area of training. This may be due to teachers’ tendency to rely on the ready-made tests found in the textbook.

On the purpose of testing scale, teachers expressed an urgent need for training in “finding out what needs to be taught/learned.” This is in contrast with Vogt and Tsagari (2014), who found that teachers desired more training in “awarding certificates.” Surprisingly, it appears that “giving grades” is not important to Saudi EFL teachers, as they did not wish to receive training in this area. This finding is consistent with those of Hasselgreen et al. (2004) and Vogt and Tsagari, who also discovered that the majority of teachers did not wish to receive training in this area. Additionally, the answers to the open-response items supported that the majority of teachers did not wish to receive training on the “purpose of testing.”

Regarding the content and concept of the language assessment scale, teachers reported a strong need for training in “testing/assessing integrated language skills.” This finding agrees with that of Vogt and Tsagari (2014), who found that assessing integrated skills was the most desired area for training. This may be because assessing more than one skill is harder than assessing one skill. However, teachers did not wish to receive future training in assessing receptive skills.

Implications

The current study revealed that Saudi EFL teachers take their role in language assessment seriously and are willing to participate in additional training programs to develop their assessment literacy, with varying priorities depending on the context. Therefore, in this section, suggestions and implications are presented to assist EFL teachers in this area. First, teachers repeatedly noted that they have not yet achieved the required level of assessment literacy. Based on this finding, the present study suggests providing professional development opportunities for EFL teachers to meet their assessment needs. This could take the form of regular workshops that would contribute to the development of LAL. Second, teachers and decision makers should work collaboratively to develop assessment training programs (Stiggins, 1999). Third, formal language assessment courses at the BA and MA levels can help teachers reach the required level of competence in LAL. However, such courses should consider the instructors’ prior experience and practices as well as the findings of LAL needs analyses like the ones described in this study. Finally, since the EFL teachers in this study have not been trained on integrating technology into their assessment practices, future assessment training programs should incorporate the use of technology.

Conclusion, Limitations, and Further Recommendations

To conclude, this study set out to investigate Saudi EFL teachers’ assessment literacy levels and language assessment training needs. The study consisted of a quantitative questionnaire that was administered to female EFL teachers in the Saudi context. Regarding the first research question, it can be said that Saudi EFL teachers lack a well-developed knowledge of assessment literacy. Questionnaire results showed that the majority of participants received either no training or basic training in LAL. For the second research question, the findings showed that there is a need for training with priority being given to different content areas of LTA.

Despite these research findings, some limitations must be taken into account. This study was limited to female EFL instructors at King Abdulaziz University because it was difficult to

access male participants' data. Additionally, the use of the questionnaire yielded only a partial picture of training levels and needs in LTA. Therefore, the use of varied instruments, such as interviews, would provide a more comprehensive view of teachers' assessment literacy. Another limitation is that the sample size was not large enough to generalize the findings to other populations.

Further research on LAL in the Saudi context is recommended, as few studies have been conducted on this topic. Additional areas of assessment literacy should be explored using other samples, such as male teachers with different educational levels. Another suggestion is to study the effects of previous training on the current assessment practices of EFL teachers. Finally, in order to obtain comprehensive results, using mixed methods, such as interviews, together with the questionnaire is highly encouraged.

About the Authors

Faten Ali Salami is a Master student in TESOL at King Abdulaziz University. She has published several research articles in peer reviewed international journals. Her research interests include students' engagement in EFL classes, Feedback in L2 writing, and language assessment. ORCID ID: <https://orcid.org/0000-0002-5751-4622>

Rasha Mastour Alharthi currently does her Master in TESOL at King Abdulaziz University. Rasha previously taught English in the English Language Center (ELC) at Taif University. She has a CELTA certificate in teaching English. Her research interests include teacher professional development and language assessment. ORCID ID: <https://orcid.org/0000-0003-1435-4726>

References

- Berry, V., Sheehan, S., & Munro, S. (2019). What does language assessment literacy mean to teachers? *ELT Journal*, 73(2), 113–123. <https://doi.org/10.1093/elt/ccy055>
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research method in education*. Routledge.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347.
- DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, 17(4), 419–438.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.
- Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Corwin Press.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132.
- Hakim, B. (2015). English language teachers' ideology of ELT assessment literacy. *International Journal of Education and Literacy Studies*, 3(4), 42–48. <https://doi.org/10.7575/aiac.ijels.v.3n.4p.42>

- Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp.413–427). De Gruyter.
- Hasselgreen, A., Carlsen, C., & Helness, H. (2004). *European survey of language testing and assessment needs. Part one: General findings*.
<http://www.ealta.eu.org/documents/resources/survey-report-pt1.pdf>.
- Huang, J., & He, Z. L. (2016). Exploring assessment literacy. *Higher Education of Social Science*, 11(2), 18–27. <https://doi.org/10.3968/8727>
- 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.
<https://doi.org/10.1177/1362168816684366>
- Koh, K. H. (2011). Improving teachers' assessment literacy through professional development. *Teaching Education*, 22(3), 255–276. <https://doi.org/10.1080/10476210.2011.593164>
- Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the language assessment literacy survey. *Language Assessment Quarterly*, 17(1), 100-120.
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169–197.
- Lan, C., & Fan, S. (2019). Developing classroom-based language assessment literacy for in-service EFL teachers: the gaps. *Studies in Educational Evaluation*, 61, 112–122.
- Mendoza, L. A., & Arandia, B. R. (2009). Language testing in Colombia: A call for more teacher education and teacher training in language assessment. *Profile: Issues in Teachers' Professional Development*, 11(2), 55–70.
- Muoz, 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.
- Newfields, T. (2006). Suggested answers for assessment literacy self-study quiz #1. Shiken: *JALT Testing & Evaluation SIG Newsletter*, 10(2), 25–32.
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381–402.
- Qassim, J. A. S. (2008). *Teachers' perceptions of current assessment practices in public secondary schools in the state of Qatar* [Unpublished doctoral dissertation]. University of Hull, Qatar.
- Shim, K. N. (2009). *An investigation into teachers' perceptions of classroom-based assessment of English as a foreign language in Korean primary education* [Unpublished doctoral dissertation]. University of Exeter, Exeter.
- Stiggins, R. J. (1991). Relevant classroom assessment training for teachers. *Educational Measurement: Issues and Practice*, 10(1), 7–12.
<https://doi.org/10.1111/j.1745-3992.1991.tb00171.x>
- Stiggins, R. J. (1995). Assessment literacy for the 21st century. *The Phi Delta Kappan*, 77(3), 238–245.
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758–765.

Umer, M., Zakaria, M. H., & Alshara, M. A. (2018). Investigating Saudi university EFL teachers' assessment literacy: Theory and practice. *International Journal of English Linguistics*, 8(3), 345–356. <https://doi.org/10.5539/ijel.v8n3p345>

Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374–402.

Appendices

Appendix A (Ethical Approval Form)

Ethics Approval Form 2021	
<p>This application form is to be used by faculty members and postgraduate students seeking ethical approval for an individual research project where primary research focus is in the English Language Institute at King Abdulaziz University. Research must NOT begin until approval has been received from the ELI Ethics Committee.</p>	
1. Applicant Details	
1.1 Primary applicant	Rasha Alharthi – Faten Salami
1.2 Affiliation	the English Language Institute at King Abdulaziz University.
1.3 KAU ID (if applicable)	2100121 – 2100128
1.4 KAU Email (if applicable)	rmesferalharthi@stu.kau.edu.sa falisalami@stu.kau.edu.sa
1.5 Applicant Type	<input type="checkbox"/> Faculty member <input checked="" type="checkbox"/> ELI Postgraduate student <input type="checkbox"/> KAU postgraduate student <input type="checkbox"/> Other, please specify:.....
1.6 List the names and affiliation of others involved in the research group	Faten Salami Rasha Alharthi Dr. Joman Shami
2. Research Study Details	
2.1 Title	Improving Classroom-Based Language Assessment Literacy for In-Service Saudi EFL Teachers
2.2 Proposed start and end date of data collection	Start Date: 15/ 11/ 2021 End Date: 05/ 12/ 2021
2.3 Main supervisor (postgraduate students only)	Name: Dr. Joman Shami Email: jshami@kau.edu.sa
2.4 Project type: Please choose the predominant data collection research instrument(s)	<input checked="" type="checkbox"/> Questionnaire/ Survey <input type="checkbox"/> Interviews <input type="checkbox"/> Focus groups <input type="checkbox"/> Experiment <input type="checkbox"/> Classroom Observation <input type="checkbox"/> Other, please specify.....



2.5 Briefly describe the aims of this research study. Any technical terms or discipline-specific phrases should be explained. Please use no more than 300 words.	
This study aims to investigate the in-service EFL teachers' current levels of classroom-based language assessment literacy (CBLAL). Moreover, it aims to investigate the knowledge and aspects of CBLAL that they wish to improve.	
2.6 Please provide an outline of the proposed research methods and procedure. Please use no more than 300 words.	
To collect the data, online survey will be adopted and modified based on the purposes of the present research study. Spss Version 23 will be applied to analyze the data	

3. Participant Information

3.1 Who are the participants?	Female EFL instructor in the ELI
3.2 Number of participants required with justification.	The minimum is 50 participants. The justification is that we plan to publish.
3.3 How will they be recruited?	Random sampling
3.4 What is the inclusion/exclusion criteria?	Participants must be in-service instructors.
3.5 Are there any possible risks that participants may experience while participating in the research?	No
3.6 Explain whether participants will be provided with any financial incentive or other "reward" that will be used during the research	No



Declaration

I certify that the information contained in this application is accurate. I have attempted to identify the risks that may arise in conducting this research and acknowledge my obligations and the rights of the participants. I confirm that the research will be conducted in line with KAU ethical code, will avoid conflict of interest, and will preserve the confidentiality and dignity of the participants.

Name of primary applicant: Rasha Alharthi & Faten Salami
 Signature: Rasha Alharthi & Faten Salami
 Date: 05/11/2021

Approval

The ELI Ethics Committee has considered the ethical aspects of this application. The committee recommends that the project be:

Approved deferred (for reasons attached) not approved

.....

Dr Badiah Hakeem ELI Vice-Dean women's campus		.18/ .11/ 2021
Dr Nashwa Saaty Head of the Graduate Studies and Academic Research Unit		.17/ 11/ 2021
Dr Hana Jan Head of the Academic Affairs Unit		.18/ .11/ 2021

Appendix B (The Questionnaire)

Part one: Background information

1. Age
 - 25 to 30
 - 30 to 40
 - 40 to 50
 - 50-60
2. Educational level
 - Bachelor
 - Master
 - Doctoral
- 3- During your pre-service or in-service training, have you learned something about testing and assessment (theory and practice)
 - Yes
 - No

Please specify if you received or need training in the following domains based on the three choices (No, yes, basic training, yes advanced training). Each domine was divided into two aspects the first one for the received training and the second for the needed training.

Part two: questions about training in language testing and assessment (LTA)

1. Preparing classroom tests

2. Using ready-made tests from textbook packages or from other sources
3. Giving feedback to students based on information from tests/assessment
4. Using self- or peer-assessment
5. Using an informal, continuous, non-test type of assessment
6. Using the European Language Portfolio, an adaptation of it or some other portfolio

Part three: questions about training in Purposes of testing

1. Giving grades
2. Finding out what needs to be taught/ learned
3. Placing students into courses, programs, etc.
4. Awarding final certificates (from school/program; local, regional or national level)

Part four: questions about training in Content and Concept of Language Testing and Assessment

1. Testing/assessing of receptive skills (reading/listening)
2. Testing/assessing of productive skills (speaking/ writing)
3. Testing/assessing of micro-linguistic aspects (grammar/vocabulary)
4. Testing/assessing of integrated language skills

Part Five: Questions to explore teachers' perceptions of their language testing and assessment competencies and training.

1. During your pre-service training, what did you learn about language testing and assessment?
2. Have you received in-service training in Language Assessment and Testing? If yes, what was the focus of this training?
3. Have you received training to use technology in language assessment and testing? if yes, what was the focus of this training?
4. Which aspects of Language Assessment Literacy (LAL) do you need to have training on and why?
 - a) Purpose of testing
 - b) Preparing and administering language tests
 - c) Scoring and rating
 - d) Giving feedback to students based on the information from tests
 - e) Assessment policy and local practices.