

RESEARCH

Open Access

EFL teachers' perceptions and practices of their language assessment knowledge



Hossein Farhady¹ and Kobra Tavassoli^{2*} 

* Correspondence: kobra.tavassoli@kiau.ac.ir

²Department of ELT, Karaj Branch, Islamic Azad University, Karaj, Iran
Full list of author information is available at the end of the article

Abstract

Research on language assessment knowledge (LAK) of teachers has focused on two major topics: identifying the LAK needs of teachers and developing appropriate LAK tests. Although the prior research findings significantly contributed to our understanding of the parameters of LAK, they were mostly quantitative and did not provide much information about EFL teachers' perceptions and applications of their LAK in a direct and face-to-face situation. Therefore, this qualitative study was designed to shed light on some key issues related to teachers' LAK using semi-structured interviews. The issues included EFL teachers' perception of their LAK and their utilization of LAK in their teaching. The participants were 11 teachers with a high level of LAK and 10 teachers with a low level of LAK determined by their performance on a LAK test. The interviews were recorded, transcribed, and content analyzed. The findings did not reveal significant differences in the responses provided by the two groups of teachers. Further, to investigate the extent of teachers' application of LAK in classroom contexts, some of the tests made by the participating teachers were collected and content analyzed. The results showed that teachers with high LAK wrote longer tests with more varied sections and tasks. Finally, no meaningful relationship was found between the teachers' level of LAK and their students' performance on classroom achievement tests. The findings imply that the language assessment field needs more research on multiple dimensions of LAK.

Keywords: Assessment knowledge, Language assessment knowledge, EFL teachers, Teacher-made tests, Teachers' perception, Teachers' practice, Teachers' voice

Introduction

Since Stiggins (1991) used the term assessment literacy (AL) as the minimum requirement for assessment knowledge (AK) of teachers, this concept has become a key issue in the assessment literature. AL refers to teachers' understanding of assessment procedures, their capacity of developing assessment tasks and criteria to evaluate students' performance, and taking proper actions upon the obtained information (Hay & Penney, 2013). Scholars believe that AL/AK can empower different stakeholders, especially teachers, by assisting them to make better decisions about the development, administration, and use of assessment tasks (Grabowski & Dakin, 2014; Popham, 2009).

Following this trend in general education, the concept of AL/AK gained significance in the language assessment field, too. Allocating the special issues of *Language Testing*

Journal (2013) and *Papers in Language Testing and Assessment* (2017), as well as the 39th Language Testing Research Colloquium (LTRC) in Colombia-Bogotá (2017) to the topic of language assessment literacy (LAL), is a witness of the significance of the issue in the language assessment field.

Although many EFL stakeholders need to improve their LAL/LAK, language teachers seem to be the most targeted group since they are the acting agents of testing and assessment procedures in the real context of education (Scarino, 2013). One of the main reasons for the necessity of LAK for teachers is that regardless of any content knowledge expected from teachers in any field, they all need a reasonable knowledge of assessment to evaluate students' achievement (Farhady & Tavassoli, 2018; Tavassoli & Farhady, 2018). Teachers who perform assessments appropriately can identify their students' needs, monitor their learning, alleviate their problems, and enhance their achievement. Further, a reasonable command of LAK can help teachers design and administer decent tests, and collect, analyze, and interpret assessment-related data to make fair decisions (Inbar-Lourie, 2013). As Berry et al. (2019) stated, effective assessment can promote learning; hence, a language teacher needs to be able to engage with a range of teaching, learning, and assessment practices simultaneously.

Despite the importance of LAK, however, research findings indicate that many teachers do not possess sufficient LAK and feel that they need help in making assessment-related decisions (Mertler & Campbell, 2005). Furthermore, a considerably rich literature on LAK demonstrates that the majority of EFL teachers do not seem to have an adequate command of language assessment issues (Baker & Riches, 2018; Berry et al., 2019; Deygers & Malone, 2019; Farhady & Tavassoli, 2015, 2017, 2018; Firoozi et al., 2019; Giraldo, 2018; Inbar-Lourie, 2008; Janatifar & Marandi, 2018; Koh et al., 2018; Kremmel & Harding, 2020; Malone, 2008, 2013; Ölmezer-Öztürk & Aydin, 2018, 2019; Tajeddin et al., 2018; Taylor, 2013; Yastıbaşa & Takkaç, 2018). A common issue across the findings of these research reports highlights certain gaps between what teachers claim about their knowledge and application of their knowledge on the one hand and what they do in the real context of instruction on the other hand.

One reason for such a gap may be the fact that the majority of the abovementioned studies were quantitative in nature and indirect in their data collection processes. The existing literature seems weak in providing information obtained through using qualitative approaches in a face-to-face context. To fill the gap, therefore, the present study addressed these issues to echo EFL teachers' voices about their perception of what they think of LAK, what they do in their test development processes, and the relationship between their LAK level and their students' achievement. The findings of this study would contribute to the field by adding data to the previous research and cross-checking the findings of quantitative and qualitative methods.

Review of the related literature

Assessment knowledge (AK) and language assessment knowledge (LAK)

As early as the 1990s, the American Federation of Teachers, the National Council on Measurement in Education, and the National Education Association (AFT, NCME, & NEA, (1990) developed some guidelines and standards to help teacher educators and teachers recognize the significance of LAK in and out of classroom context. The guidelines focused on identifying and designing appropriate assessment methods, scoring

and interpreting the results of various assessment techniques, using the results of the assessment in making decisions about students and programs, developing appropriate grading systems, and reporting assessment results to different stakeholders. Later, Stiggins (1991) reformulated the guidelines under the cover term of “assessment literacy” as the basic level of “assessment knowledge” in general education. According to Stiggins (1995), assessment literate teachers know “what they are assessing, why they are doing so, how best to assess the achievement of interest, how to generate sound samples of performance, what can go wrong, and how to prevent those problems before they occur” (p. 240). A decade later, Webb (2002) proposed a comprehensive definition of AL/AK with three main elements as having the knowledge of how to assess students, how to interpret assessment results, and how to enhance student learning and program efficiency based on assessment results.

Further, Popham (2004, 2011) capitalized on the vital role of teachers’ assessment literacy in students’ learning and stated that AL/AK should include one’s understanding of the essential issues and procedures in an assessment that influence educational decisions. He believed that many teachers know little about assessment (Popham, 2009). He also mentioned that nowadays, teachers must develop a level of understanding and knowledge in assessment since the appropriate use of assessment is a powerful tool for students’ learning (Popham, 2011). However, achieving AL/AK is not possible with a brief mention of assessment in a teacher education course. Rather, teachers need to attend a comprehensive course in assessment to acquire the necessary knowledge about what assessment is and how to implement it in the classroom. Despite all the emphasis, most studies have indicated that teachers are short of the necessary knowledge and skills on how to assess their students’ achievement effectively. For example, Razavipour et al. (2011) found that teachers generally suffered from a poor knowledge base in assessment and they resorted to external tests to evaluate their students’ achievement. Further, the results of Mede and Atay’s (2017) study showed that teachers had limited assessment knowledge and they needed training in different areas of assessment, especially in assessing the four skills in classroom contexts. Along the same line, Tsagari and Vogt (2017) found that since teachers did not have sufficient knowledge to implement the required assessment activities in the classroom, they tended to resort to traditional assessment procedures.

Parallel with the emergence of the concept of AL/AK in general education, it also entered the language education field. Language assessment literacy (LAL) as the basic level of LAK requires different stakeholders in language education to have a command of assessment-related issues (Malone, 2008). According to Inbar-Lourie (2008), a language assessment literate individual needs to know the why (the reasoning for assessment), the what (the description of the construct to be assessed), and the how (the assessment process) of assessment. Fulcher (2012) described LAL/LAK as having the knowledge, abilities, and skills needed to design both standardized tests as well as classroom tests. He also added that teachers need to be familiar with test processes and be aware of the psychometric, social, and ethical principles underlying assessment. Further, Malone (2013) defined LAL as “language instructors’ familiarity with testing definitions and the application of this knowledge to classroom practices in general and specifically to issues related to assessing language” (p. 329). In short, LAL/LAK that requires the combination of general assessment knowledge and language-specific

assessment knowledge refers to teachers' understanding of the principles underlying selecting or designing tasks, evaluating student work, and using assessment to support student learning of a new language (Inbar-Lourie, 2017; Koh et al., 2018).

Because of the importance of the concept of LAK, scholars have suggested different models of LAK in the language assessment literature. Stabler-Havener (2018) summarized some of these models mentioned below. One of the early models of LAK, introduced by Brindley (2001), includes components such as the social context of assessment, defining proficiency, constructing and evaluating language tests, assessment in the language curriculum, and putting the assessment into practice. Later, Inbar-Lourie (2008) introduced another model of LAK which dealt with the why, what, and how of assessment. The why component referred to the reason for the assessment, the what component dealt with the trait being assessed, and the how component referred to the method of assessment. Davies (2008) introduced another popular LAK model focusing on three components of skills, knowledge, and principles. The skills component dealt with item writing, test analysis, using statistics, providing reports, and the ability to use software programs. The knowledge component referred to the background in measurement, description of the language, and contextualization of language assessment, and the principles component dealt with language test use, impact, ethics, and professionalism. Further, Fulcher (2012) provided a practices, principles, and contexts model of LAK. The practices dimension was concerned with the practice of language testing including the knowledge, skills, and abilities. The principles dimension was related to the guidance for practice including processes, principles, and concepts; and the contexts dimension was about the origins, reasons, and impacts of historical, social, political, and philosophical frameworks. Pill and Harding (2013) introduced a different model of LAK, taking into account five different components of illiteracy (not knowing the language assessment concepts and methods), nominal literacy (understanding that a specific term relates to assessment), functional literacy (having a sound understanding of basic terms and concepts), procedural and conceptual literacy (understanding the central concepts and using knowledge and practice appropriately), and multidimensional literacy (knowing ordinary concepts). Finally, a LAK model was introduced by Taylor (2013), which included eight dimensions of knowledge of theory, technical skills, principles and concepts, language pedagogy, sociocultural values, local practices, personal beliefs/attitudes, and scores and decision-making. The variety of the models of LAK within a decade shows how important this concept is in the language assessment field.

Even though the concept of LAK in different models has been mostly associated with language testers (Malone, 2013), there is an urgent need to develop EFL teachers' LAK (Tsagari & Vogt, 2017) as well. The reason for the growing interest in developing teachers' LAK in the language assessment field is mainly due to the significant assessment responsibilities that new developments in language education have put on teachers' shoulders. Teachers are one of the major groups of stakeholders dealing with LAK since their assessment knowledge plays a crucial role in their assessment quality of student achievement (Koh et al., 2018; Popham, 2011).

Literature review indicates that EFL teachers' LAK has been the focus of investigations in recent years and in different places around the world (Boraie, 2015; Jannati, 2015; Mellati & Khademi, 2018; Prasetyo, 2018; Sultana, 2019; Vogt et al., 2020), to

name a few. However, much of the discussion on teachers' LAK has roots in experts' opinions of what EFL teachers know, do not know, or should know about language assessment issues (Berry et al., 2019). Teachers' direct involvement in these issues has not played a significant role in the data collections processes. To bridge the gap, this paper reports on the findings of face-to-face interviews with teachers on what they think about LAK, their knowledge about LAK, how they utilize their LAK in the classroom, how they apply their LAK in their tests, and the relationship between their LAK level and their students' achievement.

Empirical studies on EFL teachers' LAK

Researchers in different parts of the world have investigated various aspects of EFL teachers' LAK in recent years. Janatifar and Marandi (2018) explored the components of LAK and reported three components including test design and development, large-scale standardized and classroom assessment, and reliability and validity. The findings revealed that in addition to the theoretical issues of assessment, EFL teachers should also receive hands-on skills-based instruction in language assessment in their training sessions. In addition, following their findings from research on teachers' LAK, Mehmamdoost et al. (2019) recommended that EFL teachers' LAK development should receive more emphasis in teacher education programs to improve the quality of their assessment activities that may, in turn, result in better students' achievement.

Another line of research on LAK has led to the development of various LAK Scales (LAKS). For instance, Ölmezer-Öztürk and Aydın (2018) developed a LAKS after performing a comprehensive literature review to determine its content. Then they created an item pool and had the items reviewed by experts. Finally, they piloted it and reported that the test showed acceptable validity and reliability indexes. Ölmezer-Öztürk and Aydın (2019) did a follow-up qualitative study of the data collected from EFL teachers as assessors by examining their opinions and needs about language assessment. The findings revealed that the inadequacy of training in pre-service and in-service teacher education programs was the main reason underlying the insufficient LAK of EFL teachers. They recommended that the field needed more research on EFL teachers' LAK both quantitatively and qualitatively.

In another line of research, Berry et al. (2019) investigated teachers' LAK by asking them directly about their attitudes toward assessment and their assessment practices. They used interviews, classroom observations with follow-up interviews, and focus-group discussions with teachers about their training in assessment and their attitudes toward it. Although the teachers showed low confidence in their LAK, they demonstrated command of a wide range of assessment techniques. Besides, they considered the assessment activities as teaching activities and showed a negative attitude toward traditional testing and grading. Further, Firoozi et al. (2019) investigated the assessment needs of Iranian EFL teachers regarding the new assessment reform era where traditional discrete-point testing policies were dominant. They conducted in-depth interviews with EFL head-teachers and examined the documents about the curriculum reform carefully. Their findings revealed that teachers' perceptions of language testing need to change through offering them training in both knowledge and skills of language assessment. They also recommended that teachers need more training on developing

rubrics for speaking and writing as well as for developing higher-order thinking skills assessment in reading and listening comprehension. Finally, Vogt et al. (2020) studied in-service EFL teachers' beliefs about their LAK needs with the purpose of checking their perceptions of assessment, their LAK levels, and their needs. They found the context of education as one of the significant dimensions of needs analysis and proposed recommendations for context-specific teacher education programs for LAK.

In line with the qualitative approaches mentioned above, this study was an attempt to add two more sources of data to enrich previous research. That is, in addition to direct face-to-face interviews with the teachers, we collected the actual tests that they prepared and their final evaluation of their students to see if there is any meaningful relationship between these variables. More specifically, we planned to answer the following research questions.

RQ1. What do EFL teachers think about the significance of and their level of LAK?

RQ2. How do EFL teachers utilize their LAK in their teaching?

RQ3. How do EFL teachers apply their LAK to their self-developed tests?

RQ4. What is the relationship between EFL teachers' level of LAK and their students' achievement?

Method

Participants

Out of 21 EFL teachers participating in this study, 11 teachers who received scores above the mean on a LAK test formed the high group (HG), while 10 teachers who received scores below the mean on the test formed the low group (LG). The LAK test was a reliable and valid test designed to assess teachers' LAK on significant issues related to language assessment (Farhady & Tavassoli, 2018). The assumption for the selection of these teachers was that their level of LAK might lead to different perceptions of LAK, applications of LAK to their classroom assessment, and relationships with their students' achievement. We used criterion sampling of the participants (Dornyei, 2007) and selected participant teachers who met predetermined criteria on the LAK test. Further, all the teachers agreed willingly to participate in this study. Table 1 presents the demographic information of the participant teachers.

None of the characteristics of teachers, including their gender, age, academic degree, major at university, teaching experience, or teaching context, was controlled in this study, as they were not of significance to the purpose of this research. The assumption was that different demographic features did not have a significant effect on the LAK level of the teachers. Following Ölmezer-Öztürk and Aydın (2019), who claimed that the demographic information of the teachers did not play a role in the outcomes of LAK-related research, we did not control factors such as gender, age, or major field of study, for either teachers or students. Participant teachers had differing teaching experiences and taught at different ages and levels of students. However, this does not mean that these factors do not have any relationship with teachers' LAK. In future studies, researchers may want to control any of these variables to examine their potential role in teachers' LAK.

Table 1 Demographic information of the participant teachers

Variable	Description
HG/LG	11 in HG, 10 in LG
Gender	12 female, 9 male
Age	21 to 55 years
Degree	16 BA, 5 MA
Major	5 TEFL, 6 English literature, 7 English translation, 3 other majors
Teaching experience	1 to 27 years
Teaching context	1 at primary schools, 4 at high schools, 16 at private institutes

Data collection sources and instruments

To answer the research questions, the data were collected from three different sources. The primary source of data for this research was semi-structured interviews that attempted to elicit information on different key issues in language assessment from EFL teachers. The interviews focused on the participants' perceptions about the significance of LAK, their perceived level of LAK, and the extent to which they could utilize their LAK in their classroom contexts. The questions in the interview protocol addressed the following issues: which assessment topics they wished to observe in language assessment books, how they wanted to expand their LAK, which areas of language assessment they wanted to improve, how they defined LAK, what they thought about LAK and their level of LAK, which skills/components they assessed and how they scored speaking and writing, how they designed a test and which types of tasks they used in their tests, what type and varieties of tests and assessment they used more, did they use statistics to examine the item characteristics, reliability, and validity of their tests, did they prepare students for standardized tests, how did they utilize assessment results to make decisions on their students' performance, and how they would change their assessment practices if they could.

The interviews lasted between 20-40 min depending on the amount of information provided by the participants. The interviews were conducted individually in person, audio-recorded, and later transcribed for content analysis using QSR NVIVO version 10.

The secondary source of data collection was sample tests developed by the participant teachers for their classroom assessment. The teachers were asked to provide some samples of the self-made tests they used in their classes in a recent semester. The tests were quizzes, midterm exams, or final exams. These tests were either achievement tests (in the case of midterm and final exams) to check the students' progress throughout the course, or diagnostic tests (in the case of quizzes) to investigate the students' strong and weak points to provide remedies for their shortcomings. The purpose was both to examine the structure and content of the tests made by the two groups of teachers and to explore the extent to which teachers with high vs. low LAK could, indeed, apply their LAK in the real tests.

The third source of data collection was the final scores of the participant teachers' students. We asked teachers to provide us with lists of the final scores of their students in a recent semester. All the scores were converted to a scale of 100 and the mean score of each teacher's list was calculated. Then, the mean scores from the two groups of teachers were compared with each other. The purpose was to investigate whether

the teachers' high vs. low LAK had any relationship with the performance of their students.

Context of the study

Although the present study was conducted in the EFL context of Iran, the similarities of the findings of the earlier phases of this project to those of other contexts seem convincing enough to assume that the reported problems and findings may apply to other contexts as well. In the first phase of this project, EFL teachers' LAK needs were identified through administering a modified and revalidated version of Fulcher's (2012) needs assessment questionnaire. The findings from administering the questionnaire to 246 EFL teachers revealed that teachers considered general topics in language assessment as essential to their profession. In addition, almost all participants stated that they needed to improve their LAK levels (Tavassoli & Farhady, 2018). Further, following the identification of the needs, a content analysis of the testing books used at university testing courses was performed to assure the correspondence between the claimed needs and the topics in the textbooks. Crosschecking the needs and topics covered in the textbooks, the authors developed a scenario-based test of LAK. The test included six scenarios focusing on significant issues extracted from both the needs analysis of the participants and the content analysis of the textbooks. The test went through multiple review stages by experts and was piloted with 50 EFL teachers. After assuring high reliability and desirable construct validity in the piloting stages, 164 EFL teachers sat for the test (Farhady & Tavassoli, 2018). The findings showed that in contrast to their claims in needs assessment, the majority of the EFL teachers had low levels of LAK. Similar to other studies e.g., (Mertler & Campbell, 2005; Razavipour et al., 2011), these results showed a significant difference between teachers' claim and their actual LAK though they showed a willingness to improve. However, although the quantitative approach utilized in many studies in the field offers useful information about these issues, they have been mostly based on indirectly obtained information. Not many studies except for (Berry et al., 2019; Firoozi et al., 2019; Janatifar & Marandi, 2018; Ölmezer-Öztürk & Aydin, 2019; Vogt et al., 2020; Yastıbaşa & Takkaç, 2018) have used direct elicitation processes to obtain information on EFL teachers' perceptions on what they think of LAK and their level of LAK, what they really do in assessing their students' language achievement, and the relationship between teachers' LAK level and their students' achievement. This was the basic motivation for the present study to address these issues through face-to-face interviews. The details of the findings and their implications and applications are discussed below.

Results

Findings for RQ1 and RQ2: Analysis of semi-structured interviews

One of the primary purposes of this research was to collect information from participant teachers in a direct way through semi-structured interviews. We transcribed all the interviews and uploaded the interview transcriptions into QSR NVIVO program version 10 in two separate files: one for teachers with high LAK scores and another for those with low LAK scores for comparison purposes. The transcripts of the interviews

with both groups were content analyzed. As Nartey (2013) stated, content analysis is “a key methodological apparatus that enables researchers to understand the process and character of social life and to arrive at a meaning, and it facilitates the understanding of the types, characteristics, and organizational aspects of documents as social products in their own right as well as what they claim” (p. 122). In the process of content analysis, we examined the transcripts carefully and identified and classified themes related to each interview question into certain categories. Another qualified researcher followed the same process on three randomly selected interview transcripts from each group of teachers and the comparison of the outcomes of the two coders showed a high degree of agreement.

Details of the teachers’ responses related to each research question are grouped and presented below. It is worth mentioning that we only reported the most repeated common and/or unique themes from the two groups of teachers’ responses. Sample responses from teachers in both HG and LG are also provided. To observe the authenticity of the responses, the teachers’ answers are not polished and they may contain some ungrammatical points.

RQ1. What do EFL teachers think about the significance of and their level of LAK?

To find answers to the first research question, we asked a few questions from the participant teachers. The questions centered around a few issues such as the kind of topics teachers would like to see in language assessment books, how they would like to improve their LAK if they felt they needed to do so, which areas of language assessment they wished to improve, what they thought about the definition and significance of LAK, and what they thought about their level of LAK. The themes extracted from the two groups of teachers’ responses (HG vs. LG) were similar for some of these questions. For example, regardless of their LAK level, teachers from both groups demonstrated a good familiarity with the basic topics in language assessment. Analysis of the themes extracted from the related interview question revealed that most of the teachers mentioned the following topics as important which need to have a place in language assessment books: steps in test design, reliability, validity, testing theories, different types of tests, tasks, and classroom assessment.

Further, teachers in both groups explained that they were willing to improve their LAK by attending workshops, in-service training classes, and studying recommended books. One of the teachers in the LG mentioned:

LG: Workshops are good because they are more practical, but books and academic classes are theoretical.

Similarly, regarding the areas of language assessment they showed interest to improve, although some common themes emerged from the teachers’ responses in both groups, there were some differences between their responses probably due to their different needs in improving their LAK. The common areas that the two groups mentioned were their desire to improve their ability to assess productive skills of speaking and writing with an emphasis on their rating skills. However, those in the HG showed

more interest in test development skills, while those in the LG mentioned various areas such as procedures for classroom assessment, dealing with standard tests, and emphasizing testing language components of vocabulary and grammar.

On the other hand, the teachers' perceptions about the definition of LAK differed across the two groups. No common theme could be extracted from teachers' responses in the two groups about what they thought LAK was and what an assessment knowledgeable teacher does in his/her classes. The teachers' answers in the HG demonstrated that they had more information about LAK than those in the LG. The common themes from the HG about what an assessment knowledgeable teacher does include assessing students throughout the semester, recognizing students' levels, preparing a test, and evaluating class participation. On the other hand, the common themes from the LG included knowing how to write a comprehensive test and having the knowledge to design a test. A sample response from each group of teachers follows.

HG: I think language assessment knowledge means to assess your students throughout the semester. This can be done by bringing different examples to class, by bringing examples that are liked by the students, and by teaching in a way that whenever you ask the students, they can provide the information that you want.

LG: Teachers have to develop a test that has, you know, all parts in it. I think the vocabulary part, the grammar part, etc. A complete test of all parts. You know, creativity is always a matter of importance. You know, the indirect way of testing or assessing students is much preferred.

Finally, similar themes were identified from teachers' responses in both groups to the interview questions about their self-assessment of their level of LAK. Despite slight differences in identifying their LAK level, the majority of teachers in both groups considered their LAK level inadequate. Besides, the reasons the teachers provided for their low self-assessment were quite similar in the two groups. The dominant themes regarding the teachers' reasons for their low level of LAK were their unfamiliarity with testing and assessment and not knowing how to assess practice. The following is a sample response from each group.

HG: Average. I think I'm not familiar with modern and new techniques in testing.

LG: It is average because I've studied English literature. I haven't studied assessment or this kind of stuff.

Overall, the comparison of teachers' responses from the HG and LG showed no systematic differences between the two groups' perceptions of LAK and its significance, and the reasons for their perceptions. A significant finding was probably the claims made by both groups that their LAK level was not sufficient to cope with

recent developments in assessment in general and in classroom assessment in particular.

RQ2. How do EFL teachers utilize their LAK in their teaching?

The second set of questions in the semi-structured interviews centered around EFL teachers' utilization of LAK in their teaching. The questions were as follows:

- Which skills/components did they assess in the class?
- How did they score tests of speaking/writing skills?
- How did they design a test?
- Which tasks did they use in their tests?
- What type of tests did they use more often?
- Did they use any statistical analyses to check item characteristics, reliability, and validity of their tests?
- Did they use any form of alternative assessments?
- Did they ever prepare students for standardized tests?
- How did they use assessment results in making decisions on their students' performance?
- How would they change their assessment practice if they had opportunities?

Interestingly, the teachers in both HG and LG mentioned common statements regarding the utilization of their LAK in their teaching. Similar themes appeared from the teachers' responses to the skills and components they assessed in their classes, such as all parts, vocabulary, grammar, listening, and reading.

Also, similar themes were found in how they scored speaking and writing skills tests. The common points for scoring speaking were paying attention to grammar, vocabulary, accuracy, fluency, and pronunciation, and the common points for scoring writing were considering grammar, vocabulary, organization, punctuation, and spelling. A sample response from each group is provided below.

HG: I focus on accuracy and fluency in speaking, fluency is very important. For writing, I pay attention to new words, and of course, I pay attention to the grammar parts.

LG: For speaking, the first one is their fluency. The second one is using correct phrases. For writing, I actually pay attention to punctuation, the kind of grammar, the different kinds of words, and the synonyms they use.

Furthermore, similar themes were identified from teachers' responses in HG and LG to other questions. Both groups claimed that in designing a test with different sections, they matched test items with the content of the textbooks and their students' level of language ability, and they decided on the length of the test, the number of items and sections, the skills to be tested, and their order beforehand. However, a notable point was that two of the teachers in the LG mentioned they just opened the book and wrote a test indicating that they did not have any concern about how to design a test.

Similar themes were also extracted from the two groups of teachers' responses to the types of tasks/items they used in their classes. They mentioned a variety of tasks such as fill in the blanks, multiple choice, cloze, matching, and question and answer, as well as using different test types including achievement, diagnostic, placement, and proficiency. The similar themes extracted from these questions showed that no matter what their LAK level was, most teachers did almost the same things when involved in actual testing.

Regarding using statistics, and checking reliability and validity, the major theme detected from most of the teachers' responses from either group was that they did not use statistics because they did not know how to do it. The teachers in the HG provided more diverse reasons for why they did not use statistics such as not having enough time or not being asked to do it. Also, for checking the validity of a test, teachers from both groups only referred to checking the content validity of their tests, whereas none of them referred to how they would check the reliability of a test. These answers showed the unfamiliarity of both groups of teachers with two of the most important concepts in language assessment, that is, reliability and validity. A sample response from each group is provided below.

HG: When students take a test, I check to see whether the content matches the book or not, but not in a scientific way.

LG: A test should be valid, should be related to the book, not out of the book.

About using different ways of assessing students, teachers from both groups used similar techniques. The common themes we extracted from both groups' responses were using class participation, assigning homework, keeping a record of the students' works, observation of students' progress, using interviews, and using a mixture of techniques. Sample responses from both groups are presented.

HG: I focus on class activity, class participation. How dynamic they are, how willing they are to take part in the group and peer work, and activities, and how much progress they have made from the beginning to the end of the term. I also keep a record of the students' works throughout the semester.

LG: I usually assign some tasks for them after the class according to what they are into, what they are interested in outside the class. Also, I use interviewing, and I use observation in the class.

Further, both groups provided similar responses for preparing students for high-stakes tests. They claimed that they used available sample tests, used preparation books, and taught the necessary tips and techniques to students. However, those who taught at lower-level ability classes said that they never prepared their students for standardized tests.

It was interesting to find out that there was a clear harmony between the responses of HG and LG teachers to the question of how they used assessment results. The common theme that appeared from their responses was finding the students' weak points and problems and working more on them. Unexpectedly, in addition to this common theme, the teachers in the LG provided more answers. This showed that despite their low LAK scores, the teachers in the LG were more concerned about their students' weak performance on the tests. The common themes the teachers in the LG mentioned were asking students about the reason for their poor performance, writing down the most frequent mistakes, and teaching those points again. A sample response from a teacher in the LG follows.

LG: You know, one thing that I do it all the time after each exam is I write down the most frequent mistakes in my notebook. After that, I will teach that point again on the board and I will design a very short, very small task for them based on that problem, and distribute it between them, and I ask them to complete it in a friendly group without any stress.

Finally, regarding how teachers would assess their students given the required opportunities, both groups emphasized paying attention to students' class performance and using a comprehensive test with all components and skills. Sample responses from the two groups follow.

HG: I consider both a comprehensive final exam and pay attention to class performance. To check their performance, I pay attention to everything that they do in the class.

LG: A comprehensive test, I mean making a comprehensive test is more important. And I try to use the standard tests. I mean all parts must be included, and especially the vocabulary part and grammar part are more important than other parts. I have to focus on them.

Analysis of the teachers' responses in the two groups showed more commonalities in their utilization of LAK in their teaching. Overall, the results of content analysis of the semi-structured interviews revealed that although teachers with higher LAK scores provided more comments on the assessment-related issues, both groups of teachers (HG vs. LG) mentioned more or less similar points about LAK and similar themes appeared from their responses. Further, neither group of EFL teachers in this study considered themselves knowledgeable enough about language assessment nor did they utilize their knowledge of language assessment in their teaching successfully.

Findings for RQ3: Analysis of teacher-made tests

To answer RQ3 that focused on whether and how participant EFL teachers applied their LAK in their tests, we collected some sample teacher-made tests developed by the

teachers in the two groups (HG vs. LG) to check if there was any link between teachers' LAK and the quality of their developed tests. Of course, not all teachers developed their tests. Three teachers in the HG and five teachers in the LG did not provide any tests since according to what they said they were always provided with tests developed by their test centers. The fact that some EFL teachers did not develop their tests for their students was informative since it revealed that not all teachers, regardless of the degree of their LAK, are allowed to apply their knowledge to develop tests for their students. The analysis of the tests provided by the teachers in each group showed that the tests developed by those in the HG were longer with more sections and number of task types than the ones developed by their colleagues in the LG. However, although the tests developed by the teachers in the HG included sections on grammar, vocabulary, listening, reading, and writing, none of them tested speaking. On the other hand, the tests developed by the teachers in the LG were rather short and they mostly included sections on vocabulary and grammar. Rarely did these teachers test listening, reading, writing, or speaking.

Further, the types of tasks used in these teacher-made tests were compared. The most commonly used task types by the teachers in the HG were editing, fill in the blanks, matching, multiple-choice, question and answer, sentence completion, and true/false. However, the most commonly used task types by the teachers in the LG were only fill in the blanks and sentence completion. The comparison of the tests developed by the two groups of teachers showed that the teachers in the HG were more concerned about their tests and wrote longer tests which included different sections and a wider variety of tasks in each section, while the teachers in the LG mostly limited their tests to checking vocabulary and grammar. This implies that teachers with higher LAK scores are apt to apply their knowledge in practice, especially when they get involved in developing their tests.

Findings for RQ4: Analysis of teachers' lists of students' scores

To answer RQ4 that focused on the relationship between EFL teachers' LAK level and the achievement of their students, we checked the performance of the two groups of students based on the lists of their end-of-the-semester scores provided by the participant teachers. The descriptive statistics of each group of students' scores and the related independent samples *t* test are reported in Table 2.

No significant difference was observed between the performance of the students of teachers with high vs. low LAK scores ($t = -1.76$; $p = .09$; $\alpha = .05$; $p > \alpha$). Even though the difference in the students' performance in the two groups was not significantly different, quite surprisingly, the mean score of the students of teachers in the LG (90.36) was higher than the mean score of the students of teachers in the HG (83.77). This looks strange as it was expected to see a higher mean score for the students of teachers with higher levels of LAK. One reason may be the few number of teachers (11 in the HG and 10 in the LG) whose students' scores were compared. Another probable reason can be the longer and more complex tests the teachers in the HG developed which might have caused lower students' scores or these teachers might have been stricter in giving high scores to their students. The most potential reason may be that even

teachers with high LAK could not utilize their assessment knowledge in real classroom contexts to improve their students' learning.

Discussion

This paper was a report on what EFL teachers think about LAK, how they utilize their LAK in their classes, how they apply their LAK in their tests, and the relationship between their LAK and their students' achievement. Twenty-one EFL teachers who took a LAK test (Farhady & Tavassoli, 2018) were selected based on their LAK scores (11 teachers with high LAK scores and 10 teachers with low LAK scores). The teachers in the two groups participated in semi-structured interviews. The interviews were transcribed and subjected to detailed content analysis using QSR NVIVO version 10. The findings showed that in almost all aspects attempted in the interview questions, there were more similarities than differences in the responses given by the two groups of teachers. Overall, even though the teachers in the high group were more knowledgeable in assessment-related issues, and provided more detailed answers to the interview questions, their perception of LAK, their self-assessment of their level of LAK, and their utilization of LAK in their teaching did not significantly differ from their colleagues in the low group. However, the content analysis of some sample teacher-made tests that the two groups of teachers developed and used in their classes in a recent semester revealed that teachers with higher LAK scores prepared longer tests with more sub-components and more variety of task types. Further, the comparison of the performance of students of the two groups of teachers based on their reported scores at the end of a recent semester showed no significant difference between them.

The findings also showed a slight trend that having a higher level of LAK helps EFL teachers in their classroom teaching that may somehow improve student learning. Since assessment knowledge as part of teachers' professional knowledge (Farhady & Tavassoli, 2018; Tavassoli & Farhady, 2018; Popham, 2011) is essential for successful teaching, it seems vital for supervisors, teacher educators, and policy-makers to provide ample opportunities for teachers to improve their LAK (Popham, 2009). The major reason for the low level of EFL teachers' LAK is most probably the insufficiency of teacher education courses since teachers do not receive sufficient training on assessment, and therefore, their knowledge is limited (Ölmez-Öztürk & Aydın, 2019). The results of this study were also in line with Jannati (2015) where the teachers were familiar with the basic concepts in language assessment; however, despite their knowledge in assessment, they could not reflect it in their practices. For example, they could not put their knowledge of reliability, validity, fairness, and authenticity into practice. Nonetheless, just giving information to teachers about language assessment does not seem effective enough as the results of the semi-structured interviews in this study demonstrated. Teachers need to understand and learn how to utilize their LAK in practice in their

Table 2 Descriptive statistics and the independent samples *t* test of the two groups of teachers' students' scores

Group	Mean	SD	<i>t</i> test	df	Sig. (2-tailed)
HG	83.77	8.63	-1.76	17	.09
LG	90.36	7.13			

teaching, especially in classroom contexts and in using formative assessment (Popham, 2009). This may be possible through ongoing training courses on assessment in which LAK is taught and worked on through hands-on activities with teachers by testing experts (Ölmezer-Öztürk & Aydin, 2019). These courses should be designed to improve the teachers' expertise in assessing language skills since appropriate assessment is a powerful means for learning (Popham, 2011). Simply attending training sessions may not lead to higher LAK levels. Rather, long-lasting pieces of training by professionals in language testing and assessment in which practical examples are given to teachers are desirable and effective (Ölmezer-Öztürk & Aydin, 2019). EFL teachers need to get a good command of designing appropriate tests and utilizing them in class to help their students improve their learning. Since experienced teachers do not know much about assessment, they cannot help new teachers either (Popham, 2011). Therefore, the main option that remained for teachers is to attend training sessions. If teachers are well-equipped to do the assessment and they understand that it is an important part of teaching, then, assessment can change into an enjoyable experience for both teachers and students. More importantly, it seems necessary for different groups of stakeholders, including experts, teachers, administrators, supervisors, policy-makers, students, and parents to come into play and work together to improve the present situation (Kremmel & Harding, 2020; Yan et al., 2017).

Conclusion

This study, in line with Jannati (2015), Ölmezer-Öztürk and Aydin (2019), and Vogt et al. (2020), was one of the few qualitative studies in the literature in which the teachers' voice about their LAK was investigated through collecting direct and face-to-face information about various aspects of EFL teachers' LAK. The major contribution of this study was to show that EFL teachers in general, and those with low LAK levels in particular, could not perform appropriate assessment practices in their classes. Equally important, it showed that there is a huge gap between teaching practices and assessment practices in the context of the classroom which is in contrast with the new theories of language assessment. Teachers seem to be quite unfamiliar with recent movements in the assessment such as learning-oriented assessment where instruction, learning, and assessment are integrated, and teachers are supposed to be the center of classroom assessment practices (Purpura & Turner, 2014). Since assessment is an irrefutable part of the teaching-learning process, and because teachers are the center of all assessment-related activities including writing items, constructing tests, and giving scores, EFL teachers should be knowledgeable enough in language assessment (Ölmezer-Öztürk & Aydin, 2019).

The findings of this study have some implications for teacher educators and policy-makers. The results suggest an informed reconsideration of teacher education programs. Practical professional development of teachers including having LAK should become one of the priorities of teacher educators to prepare teachers for their daily assessment practices. There is a demanding need to provide ongoing hands-on instruction to teachers on different language assessment issues (Malone, 2013). Policy-makers should also provide the facilities and contexts for doing research and improving the LAK levels of various stakeholders, especially teachers (Taylor, 2013). In addition, policy makers' LAK levels should be improved as well because research findings show that

their concerns and ideas considerably differ from those of language assessment scholars (Deygers & Malone, 2019). Since policy-makers make important decisions about all aspects of language education including language assessment, they should become familiar with new assessment approaches to be able to make informed policies.

This study was limited by the data collection techniques used and the small sample size. Further useful information can be collected from classroom observation on how teachers do classroom assessments and from focus-group interviews with teachers having different LAK levels on different language assessment issues. Using larger sample sizes may also provide more insights into the LAK realities of EFL teachers and how assessment issues are handled in various classroom contexts. Another limitation was related to the nature of the sample since it was not possible to differentiate between subgroups of teachers in terms of their gender, age, academic degree, major, teaching experience, and teaching context. Further research in which different subgroups of teachers are compared may provide illuminating results about multiple dimensions of LAK. This study was also limited to just the EFL context of Iran. Before making generalizations, it is desirable to do similar studies in various EFL and ESL contexts with diverse sub-groups of teachers.

Abbreviations

EFL: English as a foreign language; AL: Assessment literacy; AK: assessment knowledge; LAL: Language assessment literacy; LAK: Language assessment knowledge; HG: High group; LG: Low group

Acknowledgements

The authors wish to thank the EFL teachers who voluntarily participated in this study.

Authors' contributions

The authors made equal contribution to conduct this research and prepare the manuscript. Both authors approved the final manuscript.

Authors' information

Hossein Farhady is an associate professor of applied linguistics at Yeditepe University in Istanbul, Turkey. He received his MA in TESL and PhD in Applied Linguistics from University of California at Los Angeles in 1978 and 1980, respectively. His major interest is research on various aspects of language assessment. He has taught courses on language testing and research methods at the MA and PhD levels for the last four decades in several countries and presented papers in professional conferences. He has widely published on language testing, directed national and international research projects, and supervised over 50 MA theses and PhD dissertations.

Kobra Tavassoli is an assistant professor of TEFL at Islamic Azad University, Karaj, Iran. She is teaching language assessment, research methodology, and language teaching methodology courses at BA, MA, and PhD levels. Her areas of interest are language assessment and teacher education. She has published in different journals and presented in national and international conferences on these topics.

Funding

There is no funding for this research.

Availability of data and materials

Data is available for submission if necessary.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

¹ELT Department, Yeditepe University, İnönü Mah. Kayışdağı Cad., 26 Ağustos Yerleşimi, 34755 Ataşehir, Istanbul, Turkey. ²Department of ELT, Karaj Branch, Islamic Azad University, Karaj, Iran.

Received: 9 June 2021 Accepted: 26 July 2021

Published online: 05 October 2021

References

- American Federation of Teachers, National Council on Measurement in Education, and National Education Association (AFT, NCME, and NEA) (1990). Standards for teacher competence in educational assessment of students. *Educational Measurement: Issues and Practice*, 9(4), 30–32.
- Baker, B. A., & Riches, C. (2018). The development of EFL examinations in Haiti: Collaboration and language assessment literacy development. *Language Testing*, 35(4), 557–581. <https://doi.org/10.1177/0265532217716732>. <https://doi.org/10.1177/0265532217716732>.
- Berry, V., Sheehan, S., & Munro, S. (2019). What does language assessment literacy mean to teachers? *ELT Journal*, 73(2), 113–123. [doi.org/https://doi.org/10.1093/elt/ccy055](https://doi.org/10.1093/elt/ccy055).
- Boraie, D. (2015). *The construct of language assessment literacy as perceived by foreign language teachers in a specific context*. Paper presented at the 37th international LTRC Conference on From Language Testing to Language Assessment: Connecting Teaching, Learning, and Assessment. Toronto: Canada.
- Brindley, G. (2001). Language assessment and professional development. In C. Elder, A. Brown, K. Hill, N. Iwashita, T. Lumley, T. McNamara, & K. O'Loughlin (Eds.), *Experimenting with uncertainty: Essays in Honor of Alan Davies* (pp. 126–36). Cambridge University Press.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347. <https://doi.org/10.1177/0265532208090156>. <https://doi.org/10.1177/0265532208090156>.
- Deygers, B., & Malone, M. E. (2019). Language assessment literacy in university admission policy, or the dialogue that isn't. *Language Testing*, 36(3), 347–368. <https://doi.org/10.1177/0265532219826390>. <https://doi.org/10.1177/0265532219826390>.
- Dornyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.
- Farhady, H., & Tavassoli, K. (2015). *EFL teachers' professional knowledge of assessment*. Paper presented at the 37th international LTRC Conference on From Language Testing to Language Assessment: Connecting Teaching, Learning, and Assessment. Toronto, Canada.
- Farhady, H., & Tavassoli, K. (2017). *Developing a professional knowledge of language assessment test for EFL teachers*. Paper presented at the 39th international LTRC Conference on Language Assessment Literacy across Stakeholder Boundaries. Bogota, Colombia.
- Farhady, H., & Tavassoli, K. (2018). Developing a language assessment knowledge test for EFL teachers: A data-driven approach. *Iranian Journal of Language Teaching Research*, 6(3), 79–94. <https://doi.org/10.30466/ijltr.2018.120602>.
- Firooz, T., Razavipour, K., & Ahmadi, A. (2019). The language assessment literacy needs of Iranian EFL teachers with a focus on reformed assessment policies. *Language Testing in Asia*, 9(2), 1–14. <https://doi.org/10.1186/s40468-019-0078-7>.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>.
- Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Issues in Teachers' Professional Development*, 20(1), 179–195. <https://doi.org/10.15446/profile.v20n1.62089>.
- Grabowski, K. C., & Dakin, J. W. (2014). Test development literacy. In A. J. Kunnan (Ed.), *The companion to language assessment, Volume II: Approaches and development*, (pp. 869–889). Inc: John Wiley & Sons.
- Hay, P., & Penney, D. (2013). *Assessment in physical education: A sociocultural perspective*. Routledge.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402. <https://doi.org/10.1177/0265532208090158>. <https://doi.org/10.1177/0265532208090158>.
- Inbar-Lourie, O. (2013). Guest editorial to the special issue on language assessment literacy. *Language Testing*, 30(3), 301–307. <https://doi.org/10.1177/0265532213480126>. <https://doi.org/10.1177/0265532213480126>.
- Inbar-Lourie, O. (2017). Language assessment literacy. In E. Shohamy (Ed.), *Language testing and assessment (3rd ed.)* (pp. 257–270). Springer International Publishing. https://doi.org/10.1007/978-3-319-02261-1_19.
- Janatifar, M., & Marandi, S. S. (2018). Iranian EFL teachers' language assessment literacy (LAL) under an assessing lens. *Applied Research on the English Language*, 7(3), 361–382. <https://doi.org/10.22108/are.2018.108818.1221>.
- Jannati, S. (2015). ELT teachers' language assessment literacy: Perceptions and practices. *The International Journal of Research in Teacher Education*, 6(2), 26–37.
- 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>. <https://doi.org/10.1177/1362168816684366>.
- Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the language assessment literacy survey. *Language Assessment Quarterly*, 17(1), 100–120. <https://doi.org/10.1080/15434303.2019.1674855>.
- Malone, M. E. (2008). Training in language assessment. In E. Shohamy, & N. Hornberger (Eds.), *Encyclopedia of language and education, Volume 7: Language testing and assessment (2nd ed.)* (pp. 225–239). Springer Science and Business Media. https://doi.org/10.1007/978-0-387-30424-3_178.
- Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329–344. <https://doi.org/10.1177/0265532213480129>. <https://doi.org/10.1177/0265532213480129>.
- Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi*, 168(1), 1–5. https://doi.org/10.1501/Dilder_0000000237.
- Mehmandoust, Z., Alavi, S. M., & Kaivanpanah, S. M. (2019). An exploration of teachers' assessment literacy in Iran. *Foreign Language Research Journal*, 8(2), 645–663. <https://www.sid.ir/en/journal/ViewPaper.aspx?id=650676>.
- Mellati, M., & Khademi, M. (2018). Exploring teachers' assessment literacy: Impact on learners' writing achievements and implications for teacher development. *Australian Journal of Teacher Education*, 43(6), 1–18. <https://doi.org/10.14221/ajte.2018v43n6.1>.
- Mertler, C. A., & Campbell, C. (2005). *Measuring teachers' knowledge and application of classroom assessment concepts: Development of the assessment literacy inventory*. <http://files.eric.ed.gov/fulltext/ED490355.pdf>.

- Nartey, M. (2013). A speech act analysis of status updates on Facebook: The case of Ghanaian university students. *Language in India*, 13(12), 114–141.
- Ölmezer-Öztürk, E., & Aydın, B. (2018). Toward measuring language teachers' assessment knowledge: Development and validation of language assessment knowledge scale (LAKS). *Language Testing in Asia*, 8(20), 1–15. <https://doi.org/10.1186/s40468-018-0075-2>.
- Ölmezer-Öztürk, E., & Aydın, B. (2019). Voices of EFL teachers as assessors: Their opinions and needs regarding language assessment. *Eğitimde Nitel Araştırmalar Dergisi - Journal of Qualitative Research in Education*, 7(1), 373–390. <https://doi.org/10.14689/issn.2148-2624.1.7c1s.17m>.
- 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>. <https://doi.org/10.1177/0265532213480337>.
- Popham, W. J. (2004). Why assessment illiteracy is professional suicide. *Educational Leadership*, 62(1), 1–2.
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 48(1), 4–11. <https://doi.org/10.1080/00405840802577536>.
- Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator*, 46(4), 265–273. <https://doi.org/10.1080/08878730.2011.605048>.
- Prasetyo, A. H. (2018). Language assessment literacy development: A student-teacher's experiences in teaching practice program. *LLT Journal: A Journal on Language and Language Teaching*, 21(2), 199–206. <https://doi.org/10.24071/lt.2018.21.0209>.
- Purpura, J. E., & Turner, C. E. (2014, October). *A learning-oriented assessment approach to understanding the complexities of classroom-based language assessment. Paper presented at the Roundtable on learning-oriented Assessment in Language Classrooms and Large-Scale Contexts*. Teachers College: Columbia University, New York.
- Razavipour, K., Riazi, A., & Rashidi, N. (2011). On the interaction of test washback and teacher assessment literacy: The case of Iranian EFL secondary school teachers. *English Language Teaching*, 4(1), 156–161. <https://doi.org/>. <https://doi.org/10.5539/elt.v4n1p156>.
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and teacher learning. *Language Testing*, 30(3), 309–327. <https://doi.org/10.1177/0265532213480128>. <https://doi.org/10.1177/0265532213480128>.
- Stabler-Havener, M. L. (2018). Defining, conceptualizing, problematizing, and assessing language teacher assessment literacy. *Working Papers in Applied Linguistics & TESOL*, 18(1), 1–22.
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, 72(3), 534–539.
- Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238–245.
- Sultana, N. (2019). Language assessment literacy: An uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, 9(1), 1–14. <https://doi.org/10.1186/s40468-019-0077-8>.
- Tajeddin, Z., Alemi, M., & Yasaei, H. (2018). Classroom assessment literacy for speaking: Exploring novice and experienced English language teachers' knowledge and practice. *Iranian Journal of Language Teaching Research*, 6(3), 57–77. <https://doi.org/>. <https://doi.org/10.30466/ijltr.2018.120601>.
- Tavassoli, K., & Farhady, H. (2018). Assessment knowledge needs of EFL teachers. *Teaching English Language*, 12(2), 45–65. <https://doi.org/10.22132/tel.2018.74116>.
- 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>. <https://doi.org/10.1177/0265532213480338>.
- Tsagari, D., & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges, and prospects. *Papers in Language Testing and Assessment*, 6(1), 41–63.
- Vogt, K., Tsagari, D., & Spanoudis, G. (2020). What do teachers think they want? A comparative study of in-service language teachers' beliefs on LAL training needs. *Language Assessment Quarterly*, 17(4), 386–409. <https://doi.org/10.1080/15434303.2020.1781128>.
- Webb, N. L. (2002, April). *Assessment literacy in a standards-based urban education setting. Paper presented at the American Educational Research Association Annual Meeting*. Louisiana: New Orleans.
- Yan, X., Fan, J., & Zhang, C. (2017, July). *Understanding language assessment literacy profiles of different stakeholder groups in China: The importance of contextual and experiential factors*. Paper presented at the 39th international LTRC Conference on Language Assessment Literacy across Stakeholder Boundaries. Bogota: Colombia.
- Yastıbaşa, A. E., & Takkaç, M. (2018). Understanding language assessment literacy: Developing language assessments. *Journal of Language and Linguistic Studies*, 14(1), 178–193.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.