

# Exploring How ELT Teachers Perceive and Practice English Language Assessment

Ali Işık

Istinye University

Correspondence concerning this article should be addressed to Ali Işık, Istinye University, Department of English Language and Literature, Maltepe Mahallesi, Edirne Cırpıcı Yolu No:9 Zeytinburnu, Istanbul, Turkey. E-mail: [isikal@hotmail.com](mailto:isikal@hotmail.com)

As a well-designed language education program naturally requires a well-designed assessment component, the pivotal role of assessment in language education needs to be stressed. This study focuses on how English language teaching (ELT) teachers receive training in English language assessment, and how they perceive and practice assessment in Turkey. The study was conducted with the participation of 198 ELT teachers from 24 K-12 level schools and eight universities. A mixed-methods research design was chosen and the data were collected through a questionnaire, follow-up interviews, observations, personal conversations, and sample exam evaluations. The findings indicated that the assessment practices of the teachers were shaped by the teachers' language learning and teaching experiences, their intuition, adherence to assessment traditions, and the emulation of what other teachers conducted to conform to group norms. It was also observed that as the teachers did not receive proper pre-service and in-service assessment training, their assessment knowledge was low. Moreover, it was found that the assessment component of teacher training programs remained peripheral and did not help equip teachers with assessment-related theoretical knowledge and practical skills. Finally, the assessment quality in these schools was found to be low and assessment was taken as a formal requirement to grade students. In the final part of the paper, some suggestions for effective assessment are put forward.

**Keywords:** Assessment, testing, teacher training, English teaching, English language teachers

## Introduction

The demand for efficient language assessment, one of the widespread issues in language education, is on the rise. (Baker, 2016; Fulcher, 2012; Tsagari, 2016). The power of assessment on education programs has started to receive considerable attention and it is widely accepted that assessment affects teaching/learning directly and determines its quality (Menken, Hudson, & Leung, 2014; Tsagari & Banerjee, 2016; Tsagari & Vogt, 2017). Hence, the quest for valid, fair, accountable, and carefully designed assessment compatible with language teaching methodology has gained momentum (Coombe, Vafadar, & Mohebbi, 2020; Levi & Inbar-Lourie, 2020; Malone, 2013). Consequently, language assessment literacy (LAL), what language teachers should know, and how they should apply their conceptual knowledge in assessment have been studied extensively (Hill & McNamara, 2011; Leung, 2009; Rea-Dickins, 2004; Tsagari, 2016). It is argued that assessment is to be considered within the framework of language teaching methodology, not as an entity exclusive from it, and teachers are to be equipped with fundamental theoretical knowledge and practical skills and strategies to design/implement high-quality assessment (Brunfaut, 2014; Douglas, 2010; Jiang, 2017; Vogt, Tsagari, & Spanoudis, 2020). In addition to designing and implementing tailor-made assessment, they are expected to know how to utilize assessment to improve teaching by viewing it as a joint venture and by communicating assessment goals and data to stakeholders (Coombe et al., 2020). In other words, they also need to be qualified enough to exploit assessment data to evaluate a language course and its constituents and revise them both to improve student learning and meet program goals via sharing assessment data with related stakeholders (Kremmel & Harding, 2020; Norris, 2016; Scarino, 2013; Tsagari et al., 2018).

The discussions about the role of assessment in language education eventually direct attention to the quality of assessment training in pre-service and in-service teacher training programs (Coombe et al., 2020; Hill, 2017;

Purpura, 2016). ELT teacher training programs have a broad spectrum ranging from linguistics, literature, pedagogy, and guidance-counselling to approaches and methods of language teaching. With the help of this comprehensive focus, ELT teachers are expected to be equipped with the necessary knowledge and skills to teach English efficiently. However, it can be claimed that the training programs strictly emphasize language teaching methods and fail to give equal importance to assessment. Thus, the fact that a language teaching methodology adopted in a language program naturally creates its assessment system and is only complete with a sound assessment system seems to be neglected (Isik, 2020). Moreover, there prevails a common conceptualization of assessment that considers assessment as a practice following a definite period of teaching. Such a conceptualization deliberates assessment as a means to check the quality of the end-product (Barrios & Milla Lara, 2018; Tsagari et al., 2018). The research also indicates the negligence and the poor quality of language teacher training programs regarding assessment (Davies, 2008; Gan & Lam, 2020; Purpura, 2016; Spolsky, 2008; Vogt & Tsagari, 2014).

However, emphasizing issues related to language teaching methods and neglecting the assessment-related ones leaves professional teacher training programs incomplete. Such an incomplete perspective eventually impedes conceptualizing language education and assessment as the parts of a whole that complement each other (Inbar-Lourie, 2008; Levi & Inbar-Lourie, 2020; Spolsky, 2008). Furthermore, such a misconception may result in hypothesizing assessment as something secondary or something that can be handled by teachers with little or no training. Thus, the role of assessment needs to be highlighted to increase ELT teachers' awareness regarding how essential assessment is in language education. The strong focus on its role could break down the long-lasting, traditional conceptualization of assessment and create an impetus to help teachers get equipped with full-fledged theoretical and practical knowledge and skills to plan, implement, and evaluate sound, valid, and reliable assessment (Baker, 2016; Brown & Bailey, 2008; Tsagari & Vogt, 2017).

As assessment in language education has been studied extensively, it has been a major topic of interest. Ultimately, a general theoretical framework is needed to design, implement, assess, and evaluate effectively by collecting data from different language education contexts. In other words, as pointed out by Inbar-Lourie (2017), Kremmel and Harding (2020), and Vogt et al. (2020a), language teacher assessment competencies and perceived assessment needs must be regarded locally and cross-culturally to create a global body of knowledge on teacher assessment literacy, which can be exploited to design better teacher training programs in assessment. Since the evidence on assessment is scarce in Turkey, this study attempts to provide data from English as a foreign language (EFL) education in the Turkish context to contribute to evolving empirical evidence on assessment by presenting how ELT teachers perceive the notion of assessment and training on assessment. Ultimately, considering how assessment is a vital component of language teaching and education programs, this study aims to shed light on assessment by providing data from a different underexplored language education context. Moreover, this comprehensive study aims to explore how assessment is perceived by ELT teachers working at different levels of EFL education in Turkey. It is likely to contribute to redesigning the content of ELT teacher training programs to increase their assessment literacy. Ultimately, the study attempts to ameliorate the assessment practice, especially in the Turkish EFL context. It also intends to catch the attention of both teacher trainers and other authorities regarding the essential role of assessment in language education and design teacher training accordingly. Finally, it aims at contributing to data on assessment in the Turkish EFL context.

## **Literature Review**

### **Assessment Training in ELT Teacher Training Programs**

Obviously, assessment training offered in language teacher training programs is designed to help teachers perform assessment and evaluation tasks more effectively (Winterbottom et al., 2008). The content of these courses varies concerning the theoretical stance, program goals, and contextual factors (Baker, 2016; Fulcher, 2012; Levi & Inbar-Lourie, 2020; Menken et al., 2014; Purpura, 2016; Taylor, 2014; Vogt et al., 2020a). In other words, although the fundamentals of assessment are covered in the assessment courses, it is hard to talk about a fixed set of conceptual knowledge and practical skills offered to ELT teacher candidates (Brown & Bailey, 2008; Norris, 2016). Thus, the quest for forming the best possible assessment content in teacher training programs has been on-going. Foregrounded topics include assessment-related issues such as the nature and roles, the implementers, and the means and methods of assessment. (Davies, 2008; Hill, 2017; Leung, 2009; Purpura, 2016; Vogt et al., 2020a; Wall & Taylor, 2014).

Inspired by these discussions, researchers investigate English teachers' knowledge and skills in assessment. Hasselgreen, Carlsen, and Helness (2004) studied what teachers, teacher trainers, and experts in European countries knew about assessment and what they needed to know. The results indicated a low level of teacher assessment literacy. The participants lacked proper training, knowledge, and skills in all topics of investigation and needed in-service training in all of them. Vogt and Tsagari (2014) replicated the research conducted by Hasselgreen et al. (2004) and reported that the undergraduate programs did not sufficiently help the teachers develop basic concepts of assessment and practical skills to prepare and implement assessment. They tried to compensate for the insufficiency at work themselves and demanded in-service training on assessment in general. Tsagari and Vogt (2017) released the qualitative part of their study (Vogt & Tsagari, 2014) and indicated that undergraduate programs did not prepare the ELT teachers well enough in terms of assessment literacy and they felt inadequate regarding assessment and thus demanded in-service training. Other studies also attempted to explore the effects of training on language assessment practice. Baker and Riches (2018) showcased the positive impact of assessment training on language teachers' assessment literacy levels in Haitian high schools. Koh, Burke, Luke, Gong, and Tan (2018) studied how training affected the quality of assessment tasks developed by Chinese language teachers and reported that while teachers comprehended various aspects of task design, they had difficulty transferring their knowledge to practice. Vogt, Tsagari, and Csépes (2020) shared the results of a European project, Teachers' Assessment Literacy Enhancement (TALE), encompassing both EFL learners and teachers. They reported that although the teachers adopted a traditional assessment approach to testing, the learners were satisfied with it. They also pointed out low LAL levels for both the teachers and the learners as well as the need for assessment training for both groups.

### **Contextual Factors**

Besides training, researchers also indicated the impact of contextual factors on teachers' assessment practice. Yan, Zhang, and Fan (2018) highlighted the impact of contextual factors on how EFL teachers practiced assessment and perceived their training needs. They found that assessment culture, policies, expected teacher roles, and local resources and constraints affected their practice and perceived needs. Vogt et al. (2020a) also pointed out that contextual factors influenced how teachers perceived the notion of assessment, their LAL levels, and the need for assessment training. In a recent study, Levi and Inbar-Lourie (2020) noted that teachers partially benefited from what they learned in generic assessment training and mostly disregarded language-related construct components when preparing their assessment tasks. Teachers' LAL was found to be formed by generic assessment, language-specific assessment, and contextual factors. Another study underlining the importance of contextual factors was conducted by Vogt et al. (2020a) who studied ELT teachers' perceptions of LAL and their training needs in Germany and Greece. They found that although their perception of assessment was the same, the contextual factors affected how they perceived their LAL levels and training needs.

### **The Quality and Content of the Assessment Courses**

The quality and content of the assessment courses offered for foreign language teachers in universities have also captured the attention of the researchers. Bailey and Brown (1996) and Brown and Bailey (2008) investigated the content of assessment courses and compared the results across twelve years. It was found that the assessment courses covered the fundamentals of assessment, which formed the basis of assessment courses and were evolving gradually. In another study, Heritage (2007) studied the assessment training of EFL teachers in China and found that the EFL teachers were not given enough assessment training in their BA, MA, and Ph.D. programs, and received no in-service assessment training. In another study, Shohamy, Inbar-Lourie, and Poehner (2008) indicated that teachers did not receive enough quality training related to assessment and consequently they lacked fundamentals of assessment and the skills to practice quality assessment. From a different, yet similar, perspective, Xu and Liu (2009) reported that training and experience formed teachers' knowledge base, however, the way they practiced assessment was also controlled by their need to conform to group norms in a given context. Another study conducted by López Mendoza and Bernal Arandia (2009) indicated that teachers in Colombia did not receive enough quality training in assessment and they employed traditional product-oriented tests. Likewise, Jin (2010) found that the assessment courses offered to train teachers were adequate in content but the theoretical knowledge was not put into practice adequately in language classrooms. Scarino (2013) indicated that teachers needed to develop their theoretical knowledge base and be ready to implement context-sensitive assessment practices. Lam (2015) demonstrated that the language assessment training pre-service teachers received in Hong Kong was inadequate and far from bridging

the gap between theory and practice. Sultana (2019) explored the language assessment literacy levels of English teachers in Bangladesh and indicated that the assessment training the teachers received was inadequate, which subsequently prevented them from conducting effective assessment and exploit it to improve language education. In another study conducted in the Chinese context to explore the assessment training needs of university English instructors, Gan and Lam (2020) revealed that the language assessment training teachers received was not enough to meet their classroom-based assessment needs. Nevertheless, they did not want to receive advanced training in assessment due to various personal and contextual factors.

### **Assessment in the Turkish EFL Context**

Although assessment has received global attention, professional associations such as the European Association for Language Testing and Assessment (EALTA) and the International Language Testing Association (ILTA) were founded, and it is a topic studied and promoted professionally by academic journals, it has not yet aroused enough attention in Turkey. Although language testing is a mandatory course in all ELT undergraduate programs, to the knowledge of the researcher, there is only one lecturer who has a Ph.D. degree in language testing. Moreover, research on assessment is scarce in Turkey. In one of those studies, Saricoban (2011) examined the tests prepared by English teachers and pointed out validity problems especially with respect to the target grammar items. In another study, Ozdemir-Yilmazer and Ozkan (2017) found that assessment in university prep schools was dominated by the proficiency exam, which functioned as a gateway for students to continue their academic programs. In another study, Mede and Atay (2017) replicated the study carried out by Vogt and Tzagari (2014) at a Turkish university and found that the assessment literacy of the teachers was limited for classroom-based assessment and assessment-related concepts. Focusing on the student perspective on EFL assessment, the findings of the study conducted by Isik (2000) revealed students' dissatisfaction with the EFL assessment practices at all levels of EFL instruction in Turkey, which they believed to be lacking in terms of assessment. Olmezer-Ozturk and Aydin (2018) suggested the need for a language assessment literacy instrument and shared the psychometric properties of the Language Assessment Knowledge Scale (LAKS) they developed to indicate its validity and reliability. However, they did not report anything about the assessment literacy levels of teachers. Moreover, they collected data only from ELT teachers working at Turkish universities.

Despite an increase in research on language assessment, it has not received enough attention in the Turkish EFL context and subsequently, insufficient data has been accumulated to pinpoint the essential role of assessment to foster EFL education. The research is scarce and the studies carried out so far are not broad enough to provide insights into how assessment is conceptualized and practiced by ELT teachers in Turkey. Although the European-based study conducted by Vogt and Tzagari (2014) included Turkish ELT teachers, it did not report anything exclusive to the assessment practices in Turkey. Furthermore, based on the findings of the study, Isik (2020) suggested the need for comprehensive teacher training in assessment to increase the assessment literacy of ELT teachers. Thus, there is a drastic need to explore and understand assessment in EFL education, in addition to how teachers perceive and practice it in Turkey. Moreover, the data obtained from this study was meant to contribute to evolving global data on assessment from a different EFL context. In short, this study aimed to present a broader perspective about assessment because it included different types of language programs at different institutions of various levels. Thus, it helped conceptualize English teachers' assessment literacy, that is, what they knew about assessment and how they put their theoretical knowledge into practice at institutions at different levels. More specifically the study focused on the following research questions:

1. How do ELT teachers in Turkey perceive training in English language assessment?
2. How is English language assessment in Turkey perceived and practiced by the ELT teachers?

## **Methodology**

### **Participants**

English education is offered at all levels of education, primary to tertiary, in both public and private schools and institutions in Turkey. The ELT teachers in these institutions formed the target population of the study. Upon completing the bureaucratic procedures and receiving the official approval of the authorities, eight

primary schools, eight secondary schools, eight high schools, and eight universities from two major cities constituting about 30% of the population in Turkey were selected using random cluster sampling. The questionnaire was offered to all ELT teachers, 352 from K-12 schools and 157 from the universities, and among those, 125 ELT teachers from K-12 schools and 73 ELT teachers from the universities answered the questionnaire. In other words, 198 ELT teachers provided data about assessment. Two ELT teachers from each institution were selected using random cluster sampling for the follow-up interview. Using convenience sampling, one ELT teacher trainer responsible for planning and implementing in-service training participated in a personal conversation.

The data obtained from the first part of the teacher questionnaire inquiring about the teachers' demographic information revealed that 15% of the teachers had one to five years of teaching experience, 23% of the teachers had six to ten, 18% of the teachers had eleven to fifteen, 23% of the teachers had sixteen to twenty, and 22% of the teachers had 21 and above. As for the schools they teach at, 12% of the teachers were from public primary schools, 11% from private primary schools, 32% from public secondary schools, 16% from private secondary schools, and 28% of the teachers from universities. Regarding academic background, 80% of the teachers had a BA and 8% had an MA in ELT. Finally, 5% received a Certificate in English Language Teaching to Adults (CELTA), and 1% received Diploma in English Language Teaching to Adults (DELTA) training.

### **Data Collection**

Using the triangulation design type of mixed-methods research design, the data was collected with the help of the questionnaire, follow-up interview, and observation from each type of institution (public/private) at each level of education (primary, secondary, high school, and university) in the fall semester of 2018.

### **Questionnaire**

To elicit data about how the teachers perceive assessment, a questionnaire developed by Vogt and Tsagari (2014) was administered to the teachers (see Appendix 1). The teacher questionnaire was composed of two sections. The first comprised eight mixed types of items and covers the demographic information about the ELT teachers. The second consisted of three subsections including a three-point Likert-type scale about assessment-related topics. The first subsection was about classroom-based language testing and assessment (LTA) and included 11 items, the second was about the purposes of testing with eight items, and the third covered the content and concepts of LTA with 16 items. Vogt and Tsagari (2014) found that the internal consistency reliability of individual scales ranged from .80 to .93. For this particular study, the internal consistency reliability of the questionnaire was computed by using Cronbach's alpha and ranged from .78 to .93, which represents a high level of internal consistency.

### **Follow-up Interview**

To get a deeper understanding of the assessment process, two ELT teachers from each type of school at different levels of education were interviewed about the topics covered in the questionnaire. The follow-up interview included eight guiding questions exploited by Vogt and Tsagari (2014). The interviews were recorded and transcribed for analysis.

### **Observation**

The institutions were also visited once a week for one semester to observe the assessment-related practice. The purpose of the observation was to witness who prepared the assessment and how it was planned, prepared, and its data exploited. Three different teachers were observed in three different classrooms, each lasted one class hour (40 to 50 minutes). The observation was conducted using the assessment process flow chart suggested by Coombe, Folse, and Hubley (2011), which includes six major steps: planning, development, administration, analysis, feedback, and reflection. It also functions as a checklist for each step of the assessment process. A haphazardly chosen teacher-made test at the high school level was presented and evaluated using the checklist suggested by Coombe, Folse, and Hubley (2011).

### **Personal Conversation with an ELT Teacher Trainer**

To get information about the in-service teacher training in assessment that ELT teachers received, the Ministry of Education was called to request an appointment. Upon contacting the department accountable for in-service teacher training, a personal conversation (October 12, 2019) was held with one of the ELT teacher trainers.

## Data Analysis

SPSS Version 26 was used to analyse the data obtained from the questionnaires. The responses were analysed through descriptive statistics to obtain frequencies and percentages. The data elicited from the interviews were categorized and coded, and the frequencies and percentages were presented. The data obtained through the observations was also categorized and reported. The sample exam was also evaluated to give an overall idea about the exams implemented in the Turkish EFL context.

## Results

### Questionnaire

The data obtained from the teacher questionnaire were discussed in four parts, namely teacher training in assessment, classroom-focused assessment literacy, the purposes of assessment, and the content and concepts of assessment literacy.

#### *Teacher Training in Assessment*

Regarding in-service training in assessment, 4% of the teachers reported that they had received training in assessment and marked the frequency of the training as five years and above. They indicated that 1% of the training was provided by the university ELT departments, 2% by the ELT publishers, and 1% by the schools they worked in.

#### *Classroom-Focused Assessment Literacy*

Table 1 summarizes the assessment-related training the participants received in their pre-service education and the need for training in classroom-focused assessment literacy.

**Table 1**

*Classroom- focused assessment literacy and need for training*

		Training Received			Training Needed	
		Frequency	Percentage		Frequency	Percentage
Preparing classroom tests	Not at all	86	43.4	None	191	96.5
	A little	101	51.0	Basic	7	3.5
	Advanced	11	5.6	Advanced	0	0.0
Ready-made tests	Not at all	189	95.5	None	193	97.5
	A little	9	4.5	Basic	5	2.5
	Advanced	0	0.0	Advanced	0	0.0
Giving feedback based on assessment	Not at all	195	98.5	None	193	97.0
	A little	3	1.5	Basic	6	3.0
	Advanced	0	0.0	Advanced	0	0.0
Self-/Peer- assessment	Not at all	198	100	None	190	95.1
	A little	0	0.0	Basic	8	4.0
	Advanced	0	0.0	Advanced	0	0.0
Informal assessment	Not at all	198	100	None	174	87.8
	A little	0	0.0	Basic	24	12.1
	Advanced	0	0.0	Advanced	0	0.0
ELP or Portfolio	Not at all	173	87.4	None	154	77.8
	A little	25	12.6	Basic	44	22.2
	Advanced	0	0.0	Advanced	0	0.0

Table 1 indicates that an overwhelming majority of teachers had received no training in ready-made tests, giving feedback on assessment, self/peer assessment, informal assessment, and ELP or portfolio, except for *preparing classroom tests*, which was comparatively evaluated more positively in the classroom-focused assessment sub-section. As 5.6% of the teachers stated that they received advanced training on *preparing classroom tests*, about half of the teachers marked the *A little* option, and the rest (43.4%) *Not at all*. *ELP or portfolio* is the second item that relatively differed from the others. While an overwhelming majority of the

teachers picked the *Not at all* option, 12.6% of them marked the *A little option*. Although the teachers lacked adequate training on classroom-focused assessment, they did not report any significant need for training, excluding ELP or portfolio and informal assessment. Concerning training needed, 22.2% of the teachers indicated that they needed basic training in *ELP or portfolio* and 12.1% demanded basic training on *informal assessment*.

### **Purposes of Testing**

Table 2 illustrates the responses of the ELT teachers regarding the purposes of testing and the need for training in this area.

**Table 2**

*The purposes of testing and the need for training*

		Training Received			Training Needed	
		Frequency	Percentage		Frequency	Percentage
Giving grades	Not at all	194	98.0	None	196	99.0
	A little	4	2.0	Basic	2	1.0
	Advanced	0	0.0	Advanced	0	0.0
Finding out what needs to be taught/learned	Not at all	198	100	None	192	97.0
	A little	0	0.0	Basic	7	3.0
	Advanced	0	0.0	Advanced	0	0.0
Placing students	Not at all	193	97.5	None	194	98.0
	A little	5	2.5	Basic	4	2.0
	Advanced	0	0.0	Advanced	0	0.0
Awarding certificates	Not at all	195	98.5	None	198	100
	A little	3	1.5	Basic	0	0.0
	Advanced	0	0.0	Advanced	0	0.0

Table 2 shows that concerning the training received and training needed, the items in this sub-section showed a nearly homogenous distribution. The teachers did not indicate any significant training on the *purposes of testing*, nor did they mention any significant training in this topic.

### **Content and Concepts of Assessment Literacy:**

Table 3 presents the findings of the content and concepts of assessment literacy section of the questionnaire.

**Table 3**

*Content and concepts of assessment literacy and training needed*

		Training Received			Training Needed	
		Frequency	Percentage		Frequency	Percentage
Receptive skills (reading/listening)	Not at all	198	100	None	198	100
	A little	0	0.0	Basic	0	0.0
	Advanced	0	0.0	Advanced	0	0.0
Productive skills (writing/speaking)	Not at all	198	100	None	198	100
	A little	0	0.0	Basic	0	0.0
	Advanced	0	0.0	Advanced	0	0.0
Microlinguistic aspects (grammar/vocabulary)	Not at all	192	97.0	None	198	100
	A little	7	3.0	Basic	0	0.0
	Advanced	0	0.0	Advanced	0	0.0
Integrated language skills	Not at all	198	100	None	192	97.0
	A little	0	0.0	Basic	7	3.0
	Advanced	0	0.0	Advanced	0	0.0
Aspects of culture	Not at all	176	88.9	None	198	100
	A little	22	11.1	Basic	0	0.0
	Advanced	0	0.0	Advanced	0	0.0
Reliability	Not at all	198	100	None	194	98.0
	A little	0	0.0	Basic	4	2.0
	Advanced	0	0.0	Advanced	0	0.0

		Training Received			Training Needed	
		Frequency	Percentage		Frequency	Percentage
Validity	Not at all	198	100	None	198	100
	A little	0	0.0	Basic	0	0.0
	Advanced	0	0.0	Advanced	0	0.0
Using statistics	Not at all	190	95.1	None	164	83.0
	A little	8	4.0	Basic	34	17.1
	Advanced	0	0.0	Advanced	0	0.0

Table 3 illustrates the perceptions of the teachers converge on the *content and concepts of literacy* in terms of training received and training needed. Excluding microlinguistic aspects, aspects of culture, and using statistics, the teachers reported that they received no training in this sub-section. Concerning the *aspects of culture*, 11.1% of the teachers reported that they received a little training. Likewise, 4% of the teachers stated that they received a little training on *using statistics*, and 3% marked the *A little* option for *microlinguistic aspects*. Although their content and concepts of literacy was significantly low, an overwhelming majority did not demand training, except for on using statistics for which 17.1% of the teachers reported the need for basic training.

### Teacher Interviews

The data obtained from the interviews are summarized as follows:

#### Training

The data obtained from the interviews were examined concerning preservice and in-service training.

**Preservice.** All the teachers indicated that they received an assessment course. Although 92% of the teachers reported that they received an assessment course in ELT, 8% of the teachers said that they took a general course offered to all students of the faculty of education. Reporting the nature of the assessment course, 83% of those teachers said that the course was theoretical and they covered general concepts of assessment. Regarding the practical aspect of the course, 33% of them said that they prepared mostly multiple-choice type of items as practice. Evaluating the effectiveness of the course, 66% of them were not happy with the performance of the lecturer and 93% of the teachers found no relevance between the courses they received at the university and the assessment they practiced in their schools.

**In-service.** One teacher who participated in the interviews had an MA in ELT. Surprisingly she stated that she had received no courses on assessment in the MA program. Only 17% of the teachers indicated that they received in-service training on assessment and said that the content of the program was like the ones they received in university and had no practical value.

Three teachers with CELTA said that they received no training about assessment during the program. One of them reported that topics on assessment were mentioned theoretically during the training program.

#### Conceptualization of Assessment

All the ELT teachers stated that assessment was an official requirement and a tool to assess learner performance that is eventually reported in the form of grades. About one-sixth (17%) of the teachers mentioned that it was also a means of improving language programs. Moreover, about one-third (33%) of the teachers reported that the assessment created an extra burden; they would rather deal with language teaching methods and techniques, not assessment.

#### Function of Assessment

All the teachers considered assessment to be a screening process through which student performance was stratified. The exam scores were utilized to rank students and also decide who passed or failed. Besides, 8% of the teachers believed that assessment provided invaluable feedback about teacher performance and the effectiveness of teaching. Likewise, 8% of the teachers believed that it could function as leverage to revise the English program.



### ***Resources for Assessment***

With no exceptions, all the teachers said that they exploited the ready-made tests that accompanied the coursebooks prepared by international publishers, or used the tests they found on the internet. A quarter (25%) of the teachers stated that they obtained texts from the internet but prepared their test items. Only 8% of the teachers mentioned that they referred to the internet to expose themselves to alternatives for assessment and how to put them into practice.

### ***Collaboration for Assessment***

The overwhelming majority (92%) of the teachers defined collaboration as sharing the exam preparation process by dividing the number of exam tasks by the number of teachers. Then each teacher prepared the part s/he was responsible for. None of the teachers reported that they got help from their colleagues to check item quality, grammatical accuracy, or the format of the exam.

### ***Means of Assessment***

All the teachers considered the traditional pen and paper type of assessment as the common means of assessment. Additionally, 17% of the teachers said that projects affected the grades of learners. A quarter (25%) indicated homework and student participation as a means of assessment, and 8% indicated that what students did inside and outside the classroom mattered.

### ***Exploiting Assessment Data***

About one-third (34%) of the teachers said that assessment provided valuable information about student learning and it could be used to improve teaching. About one-sixth (17%) said that it provided feedback about their performance. They also implied that the administration tended to evaluate their performance in accordance with the exam scores of their students. Finally, 8% said that the data obtained from assessment could be used to evaluate the course in general.

### ***The Effects of Assessment***

All the teachers said that assessment determined the standing of learners among other students and their passing grades. All the teachers at the tertiary level reported that different types of exams served different purposes. Achievement exam scores determined if students were qualified enough to take the proficiency exam and the scores also served as a practice for their final exam.

### ***Assessment Practice***

An overwhelming (92%) percentage of teachers reported that experience, tradition, and habit shaped their assessment practices. They followed the footsteps of others and imitated the exam format adopted in their schools. The majority (83%) of the teachers felt adequate about preparing exams. None of them stated that they got help from their colleagues. Two-thirds (66%) of them stated that they preferred exams prepared by someone else, such as the testing office at the Ministry of Education.

### ***Criteria for Assessment***

All the teachers identified the target structures and vocabulary items from their coursebooks and prepared questions accordingly. They identified the items to be assessed and their weight (points allotted for each item) intuitively. At the tertiary level, the exams were developed by the testing office independent of the content of the syllabus. They served as gradual preparation for the proficiency exam. Thus, proficiency set the criteria.

### ***Need for Assessment Training***

None of the teachers mentioned any need for in-service training on assessment. Alternative assessment aroused little interest among teachers and only 8% of them reported the need for training in alternative assessment methods.

### ***Observations***

#### ***Setting***

At the schools, the lessons were carried out according to the academic plan that specified the units of the coursebooks taught in each given week. Assessment-related activities were not observed until the week before the exam was scheduled. In the meantime, one teacher from a private secondary school gave quizzes fortnightly. Two teachers from the state schools checked whether the units in the workbooks were done by the students

and pluses and minuses were given for each student and they would be added up at the end of the semester and given as an oral grade. In a private school, the department assigned the assessment tasks to each teacher from each grade two weeks before the exam date. Each teacher shared a part and then they brought their parts together to form the exam. The other private schools did the same one week before the exam. It was common practice in those schools to either gave practice tests or worksheets focusing on the topics that would be covered on the exam. In the state schools, the exam preparation procedures were the same except for the practice tests and worksheets. At the universities, before the exam week, the teachers informed students about the content of the exams in general.

### ***Feedback***

At schools, giving learners feedback about the exam took place in the form of announcing their grades. The learners had a chance to examine their exam papers to see their exam performance and how the teachers scored their exam papers. The teachers had to rush when scoring exams to deliver student scores before the parent-teacher association meeting that was held once a semester to share the exam results with the parents.

At the universities, the learners learned their grades by logging into their accounts operated by the student affairs office. They had no chance to see their exam papers unless they officially applied to examine them.

### ***Assessment Tools***

The measurements and evaluations were based on pen and paper assessment. The knowledge of grammar, vocabulary, and reading comprehension formed the basis of the exams. The exams employed decontextualized items deprived of their functional use. Discrete items that required little to no learner production mainly dominated the exams. Listening and speaking were not assessed, but writing was assessed in all schools and universities for the learners who were at and above the elementary level. Two of the universities also included listening on the proficiency exams.

### ***A Sample Exam Evaluation***

One exam from a high school was chosen randomly and evaluated. The exam was far from meeting the qualities of a well-formed, sound assessment tool (see Appendix 2). First, it did not give any information about the exam. It did not say which class it was prepared for, what topics or units it covered, the date of the exam, any general instructions, or an overall grading scheme. It was poor in terms of face validity. The teacher himself/herself did not prepare the assessment items but found a theme from different sources. S/he cut and pasted them on the exam paper. The exam paper looked like a collage, a mixture of typed and hand-written/drawn sections. The items on the exam were compressed and disorganized. The exam started with a reading text with no instructions. The second task was chart reading. Both reading texts tested grammar and the usage of the simple present tense. Moreover, the learners could have easily found clues from the previous tasks to do task B. In Part C, there were possible answers. Moreover, there were unintentional clues from Text A again, such as “have lunch, watch a film”. Part D is contextualized but only testing “am/is/are”. The learners could easily find clues from other texts including “am/is/are”. Task E tested two things at a time and lacked validity. The prerequisite task was to form the questions and then the next task was to match them with the appropriate answers. There was more than one correct answer on this task. Task F also tested two things at a time, both putting the sentences in the correct order and rewriting them considering the usage of the simple present tense. This task also provided unintentional clues for the other tasks and learners could find clues from other tasks to answer this item. Moreover, the grading was unfair because the teacher awarded two points for each item. Just writing “am, is, are” in Task D and creating the questions and matching them with the correct answers in Task E were of the same weight. Furthermore, the tasks were not organized from simple to complex. Task D was the easiest one and should have been the first task. In short, the exam employed a variety of tasks, but they were all created to test grammar, namely the correct usage of the simple present tense. To sum up, it was not meaningful, valid, or fair.

### ***The Personal Conversation with the ELT Teacher Trainer from the Ministry of Education***

To learn about the assessment-related teacher training policy and plans of the Ministry of Education, the top authority responsible for in-service teacher training, a personal conversation was held with a teacher trainer who asked that their name not be disclosed. During the conversation, it was elicited that the Ministry had no general in-service training aiming to train all the ELT teachers in certain periods. The in-service training was provided to only a few of the teachers who were nominated. No needs analysis was carried out to set the content

of in-service training sessions. The topics were determined by the teacher trainers or by the academicians invited from universities but they generally revolved around ELT methodology and techniques. Each year nearly the same in-service training with the same content was offered to ELT teachers.

The overall summary of the results revealed that the assessment training teachers received was inadequate in that they followed a traditional assessment practice either by imitating the existing widespread assessment tradition or emulating their colleagues. They tended to perceive assessment as an official procedure meant to assign grades to their students. Generally, they considered it a practice following a certain amount of teaching to test the quality of the end-product. They did not consider it to be a means of fostering language education. They were satisfied with what had they been practicing and did not mention any need for assessment training. In-service training was offered to a few teachers for a short period but it did not contribute to their assessment knowledge and skills. Put together, both the assessment training they received and the official procedures helped maintain the long-established traditional assessment practices.

## Discussion

This research attempted to explore how ELT teachers perceived the training they received on English language assessment. It also sought to explore how they perceived and practiced it in EFL education. The data obtained from the questionnaires, interviews, sample exam evaluation, and observations corroborated one another and helped answer the research questions and draw disciplined inferences about how assessment was perceived and practiced in the Turkish EFL context. The findings showed that the ELT teachers lacked proper training in assessment and perceived assessment as an official procedure to grade their students. It was not the current perspective, but their experience and long-established tradition that informed/determined their assessment practice. However, although their assessment literacy was low, they were satisfied with it and did not feel the need to improve it. Regarding the first research question investigating the assessment background of the ELT teachers, this study indicated that the training they received was insufficient and did not sufficiently cover current assessment theory and practice. It was found that the assessment courses conducted during the pre-service training were organized around theoretical issues and lacked a practicum, which meant that teachers might still not be entirely aware of the conceptualization of assessment, and therefore not be able to efficiently put it into practice. Moreover, the teachers barely received any in-service training on assessment. The findings indicated that only few teachers enjoyed the privilege of receiving in-service training. What was more surprising was that the ones with MAs in ELT reported that they took no courses on assessment in their MA programs. In a similar vein, CELTA and DELTA did not cover assessment; when they did, they only raised some basic theoretical issues. Furthermore, the interview with an ELT teacher trainer from the Ministry of Education also indicated that assessment was not a top priority. The limited in-service training for small number of ELT teachers catered to language teaching methods and neglected assessment. Thus, this finding complied with those of Gan and Lam (2020), Hasselgreen et al. (2004), Heritage (2007), Lam (2015), Scarino (2013), Shohamy et al. (2008), Tsagari and Vogt (2017), and Vogt and Tsagari (2014) who also reported that ELT teachers did not receive enough training and displayed low assessment literacy levels. In short, regarding the preservice and in-service training on assessment, the situation in Turkey was unsatisfactory, which corroborates the findings of Tsagari and Vogt (2017) and Vogt et al. (2020b). Regarding the need for training, the findings in this study were in sharp contrast with the findings in other studies mentioned above as the teachers in this study were content with their current assessment literacy and did not feel the need for assessment training. As they lacked proper training in assessment, they were not aware of their inadequacy and believed that they were assessing effectively. They perceived assessment as a formal procedure, which they performed regularly, whilst meeting the expectations of authorities to grade their students. Moreover, they did not confront any assessment-related problems and criticisms. Ultimately, an overwhelming majority of the teachers did not report any need for assessment training.

As for the second research question, it was observed that the perception and practice of assessment adopted by ELT teachers reflected a traditional understanding. The teachers at all levels of education., adopting a narrow definition and function of assessment, tended to perceive it as an official procedure to assess the performance of students at certain intervals in order to give their grades. Likewise, the teachers did not consider assessment as an ongoing evaluation of each aspect of an English course but as a vehicle to assess student performance.

They conceptualized assessment as an officially required post-teaching practice employed at the end of certain periods in the academic program. Similarly, they exploited traditional assessment techniques that verified the findings of Tsagari and Vogt (2017) and Vogt et al. (2020b). They administered pen and paper exams and considered the homework performance of the students while evaluating their performance. They also had to evaluate term projects that were required by the regulations and assigned to the students volunteering to do a project in English at the beginning of each semester. The exam dates were clearly delineated at the beginning of each academic year, which served as the assessment plan. There was no reference to basic issues such as the principles of assessment, how student performance was evaluated, what kind of assessment types -including alternatives in assessment- were employed, or how assessment was exploited to improve syllabi, teaching/learning, course materials, methodology, teacher development, and the context. At the primary, secondary (middle), and high school levels, nothing was specified about the content of the exams. Thus, administering exams on time and submitting the grades to the system might potentially pose as the main requirement(s) for fulfilling their assessment obligations.

Concerning how the teachers practiced assessment, the quality of assessment and means of exploiting the data obtained from assessment to improve programs and teaching quality did not receive the primary stress. The ELT teachers practiced assessment based on their prior language learning and teaching experience, intuition, and traditions, thus emulating what others had done so far, which was also indicated by Jin (2010) and Xu and Liu (2009). The teachers considered assessment as a kind of forced duty and were not so happy carrying it out. Before preparing the exams, they did not prepare any table of specifications or any guidelines reflecting their objectives. They preferred it to be prepared by the English Department, the testing office, or the Ministry of Education. While preparing the exams, they tended to use coursebook-accompanying tests prepared by international publishers or collected from the internet. They shared the burden of exam preparation with their colleagues and prepared their share of the items. In the private schools, assessment seemed to create frustration among the teachers. While preparing the exams and submitting the final grades of their students, they felt the need to form a fine balance between the academic success of their students and the success-oriented expectations of the administration and parents. This finding complies with the findings of Gan and Lam (2020), Koh et al. (2018), Levi and Inbar-Lourie, 2020, Vogt et al. (2020a), and Yan et al. (2018) who indicated how contextual factors affected assessment practices. At the universities, the exam dates were marked on the academic calendar and carried out accordingly. The EFL teachers were not generally responsible for exam preparation. The exams were prepared and evaluated by the testing office. This forced teachers to be success-oriented because in the report prepared by the testing office the performance of each class was analysed, and indirectly, the ELT teachers were as well. Especially in the university prep classes, all the emphasis was placed on helping the students pass the proficiency exam. Thus, at the university level, teaching for testing dominated the program and the effect of assessment was felt rigorously. This finding on assessment at the university level is in line with that of Ozdemir-Yilmazer and Ozkan (2017).

When regarded concurrently, ELT teacher training in assessment and the perception and practice of assessment showcases that the lack of proper training in assessment directly influenced how the ELT teachers conceptualized and implemented assessment. Oblivious to the current discussions on assessment, the teachers were doomed to depend on their previous EFL learning, teaching, and assessing experience, which contributed to the unproductive application of pre-established traditional testing habits. Moreover, this study showed that the lack of proper training prevented them from adopting a scientific stance as opposed to a traditional one, deeming them susceptible to contextual influence, such as the assessment perspective adopted and practiced by their colleagues. Finally, the insufficiency of training led them to adopt a narrow conceptualization of assessment and show satisfaction towards what they were practicing. In other words, the lack of proper training yielded limited assessment knowledge, which resulted in insufficient assessment practice.

## **Conclusion and Implications**

The issues raised in the purpose of the study were addressed succinctly. First of all, the results of this study are expected to contribute to the evolving global body of knowledge on teacher assessment literacy by providing empirical evidence from a different context of language education, namely Turkish EFL. Considering the local and cross-cultural need for research in language teacher assessment perceptions and competencies, the study

also offered a perspective from an underexplored EFL context. Moreover, it pioneered and attempted to portray how assessment is conceptualized and practiced from a wider perspective with the data provided by ELT teachers from different types of schools and universities as well as varying levels of education. Hopefully this encompassing study initiates a pool of local research on teacher assessment literacy to shed light on the perception of assessment practiced in the Turkish EFL context. Finally, by indicating EFL teachers' low LAL levels, the study is expected to pinpoint a mostly unexplored assessment problem, which has serious detrimental repercussions for their assessment practice. In doing so and in revealing this detrimental lack of awareness, the study is expected to make a noteworthy contribution by drawing the attention of both teachers and authorities to the fossilized assessment problem.

The study also highlighted that the teachers had low LAL levels, which gave rise to the narrow conceptualization of assessment and improper assessment practices. They cared more about preparing and administering the exam on time as indicated in their academic plan and tended to define assessment as an official requirement to evaluate what was taught. The void left by inadequate training in assessment was filled by a series of contributing factors, namely personal language learning/teaching experience, the assessment tradition at schools /universities, the need to conform to the assessment approach and practice adopted by their colleagues, and the assessment courses taken during the pre-service education. The last factor, the assessment courses conducted during the pre-service training, seemed to be the least effective because of their being too theoretical and inefficient to equip teachers with fundamental assessment literacy skills, strategies, and practical skills to prepare and evaluate relevant and sound assessment procedures. Consequently, this gap is filled with tradition, experience, imitating other colleagues, and adopting group norms.

These findings pointed out the need for pre-service and in-service training to align teaching and assessment and consider assessment as a complementary part of the teaching process. Considering the effect of assessment on language education, EFL teachers' weak assessment literacy skills could harm the language education process. However, there seems to be no authority (academic and institutional) to pinpoint the problems and propose solutions for effective assessment. Thus, by collaborating with universities, the Ministry of Education could form an assessment board to plan and implement data-driven teacher training on assessment. Consequently, an EFL assessment model that stems from the realities of that particular EFL context could be realized.

The study attempted to explore one unique aspect of the assessment-related discussions. There remain other issues to be investigated, such as: What means can be employed to help all stakeholders reconceptualize assessment? How can ELT teachers be made aware of their gaps in knowledge on current assessment practices and the need for training? How are context-sensitive pre-service and in-service training programs designed? How can practical, sustainable, career-long in-service training programs addressing all teachers be devised? Hopefully, the study can spark numerous other scholarly studies to focus on these issues to gather a rich amount of data to help make data-driven, disciplined decisions to improve the quality of assessment. Likewise, it is intended that the study contributes to making assessment in English a popular and hot topic among academics and ELT teachers and attract the attention of decision-makers. Such a focus on assessment, hopefully, ignites action to improve both assessment and language education.

This study had several limitations. The data was collected from only two major cities and failed to include all regions in Turkey. Moreover, the data obtained from public and private schools were not analyzed separately to compare these two types of schools. It would have been better to see if there were any differences between these two groups of ELT teachers concerning assessment orientation and practices.

### **List of Abbreviations**

CELTA: Certificate in English Language Teaching to Adults  
 DELTA: Diploma in English Language Teaching to Adults  
 ELT: English Language Teaching  
 EFL: English as a Foreign Language  
 LAL: Language Assessment Literacy

## References

- Bailey, K. M., & Brown, J. D. (1996). Language testing courses: What are they? In A. Cumming, & R. Berwick (Eds.), *Validation in Language Testing* (pp. 236–256). Multilingual Matters.
- Baker, B. (2016). Language assessment literacy as professional competence: The case of Canadian admissions decision makers. *Canadian Journal of Applied Linguistics/Revue Canadienne de Linguistique Appliquée*, 19(1), 63–83.
- 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>
- Barrios, E., & Milla Lara, M. D. (2018). CLIL methodology, materials and resources, and assessment in a monolingual context: an analysis of stakeholders' perceptions in Andalusia. *The Language Learning Journal*, 48(1), 60–80. <https://doi.org/10.1080/09571736.2018.1544269>
- Brown, J. D., & Bailey, K. M. (2008). Language testing courses: what are they in 2007? *Language Testing*, 25(3), 349–383. <https://doi.org/10.1177/0265532208090157>
- Brunfaut, T. (2014). A lifetime of language testing: An interview with J. Charles Alderson. *Language Assessment Quarterly*, 11(1), 103–119. <https://doi.org/10.1080/15434303.2013.869818>
- Combee, C., Folse, K., & Hubley, N. (2011). *A practical guide to assessing English language learners*. Michigan University Press.
- Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: what do we need to learn, unlearn, and relearn?. *Language Testing in Asia*, 10(1), 1–16. <https://doi.org/10.1186/s40468-020-00101-6>
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3) 327–347. <https://doi.org/10.1177/0265532208090156>
- Douglas, D. (2010). *Understanding language testing*. Hoddler Education.
- Fulcher, G. (2012) Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
- Gan, L., & Lam, R. (2020). Understanding university English instructors' assessment training needs in the Chinese context. *Language Testing in Asia*, 10(1), 1–18. <https://doi.org/10.1186/s40468-020-00109-y>
- Hasselgreen, A., Carlsen, C., & Helness, H. (2004). *European survey of language testing and assessment needs. Part one: General findings*. [www.ealta.eu.org/resources](http://www.ealta.eu.org/resources)
- Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan*, 89, 140–145. <https://doi.org/10.1177/003172170708900210>
- Hill, K. (2017). Understanding classroom-based assessment practices: A precondition for teacher assessment literacy. *Papers in Language Testing and Assessment*, 6(1), 1–17.
- Hill, K., & McNamara, T. (2011). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29(3), 395–420. <https://doi.org/10.1177/0265532211428317>
- 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>
- Inbar-Lourie O. (2017). Language assessment literacies and the language testing communities: A mid-life identity crisis? Paper presented at *39th Language Testing Research Colloquium*, Universidad de los Andes, Bogotá, Colombia.
- Isik, A. (2020). Do Students Feel that they are Assessed Properly? *Iranian Journal of Language Teaching Research*, 8(1), 63–92. <https://doi.org/10.30466/IJLTR.2020.120808>
- Jiang, J. (2017). Language assessment: critical issues—An interview with Antony John Kunnan. *Language Assessment Quarterly*, 14(1), 75–88. <https://doi.org/10.1080/15434303.2016.1269770>
- Jin, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing*, 27(4), 555–584. <https://doi.org/10.1177/0265532209351431>
- 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>
- 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>
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169–197. <https://doi.org/10.1177/0265532214554321>
- Leung, C. (2009). Developing formative teacher assessment: Knowledge, practice, and change. *Language*

- Assessment Quarterly*, 1, 19–41. [https://doi.org/10.1207/s15434311laq0101\\_3](https://doi.org/10.1207/s15434311laq0101_3)
- Levi, T., & Inbar-Lourie, O. (2020). Assessment literacy or language assessment literacy: Learning from the teachers. *Language Assessment Quarterly*, 17(2), 168–182. <https://doi.org/10.1080/15434303.2019.1692347>
- López Mendoza, A.A., & Bernal Arandia, 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.
- 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>
- Mede, E. & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi*, 168(1), 43–60. doi
- Menken, K., Hudson, T., & Leung, C. (2014). Symposium: Language assessment in standards-based education reform. *TESOL Quarterly*, 48(3), 586–614. <https://doi.org/10.1002/tesq.180>
- Norris, J. M. (2016). Language program evaluation. *The Modern Language Journal*, 100(S1), 169–189. <https://doi.org/10.1111/modl.12307>
- Olmezer-Ozturk, E., & Aydın, B. (2018). Toward measuring language teachers' assessment knowledge: Development and validation of Language Assessment Knowledge Scale (LAKS). *Language Testing in Asia*, 8(1), 1–15. <https://doi.org/10.1186/s40468-018-0075-2>
- Ozdemir-Yilmazer, M., & Ozkan, Y. (2017). Classroom assessment practices of English language instructor. *Journal of Language and Linguistic Studies*, 13(2), 324–345.
- Purpura, J. E. (2016). Second and foreign language assessment. *The Modern Language Journal*, 100, 191–208. <https://doi.org/10.1111/modl.12308>
- Rea–Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing*, 21, 249–258. <https://doi.org/10.1191/0265532204lt283ed>
- Saricoban, A. (2011). A study on the English language teachers' preparation of tests. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 41, 398–410.
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309–327. <https://doi.org/10.1177/0265532213480128>
- Shohamy, E., Inbar-Lourie, O., & Poehner, M. (2008). *Investigating assessment perceptions and practices in the advanced foreign language classroom* (Report No 1108). Center for Advanced Language Proficiency Education and Research (CALPER).
- Spolsky, B. (2008). Language testing at 25: Maturity and responsibility? *Language Testing*, 25(3), 297–305. <https://doi.org/10.1177/0265532208090153>
- 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>
- Taylor, L. (2014). General language proficiency (GLP): Reflections on the “Issues Revisited” from the perspective of a UK Examination Board. *Language Assessment Quarterly*, 11(2), 136–151. <https://doi.org/10.1080/15434303.2014.896366>
- Tsagari, D. (Ed.). (2016). *Classroom-based assessment in L2 contexts*. Cambridge Scholars Publishing.
- Tsagari, D., & Banerjee, J. (Eds.). (2016). *Handbook of second language assessment* (vol. 12). Walter de Gruyter GmbH & Co KG.
- Tsagari, D., & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment*, 6(1), 41–63.
- Tsagari, D., Vogt, K., Froelich, V., Csépes, I., Fekete, A., Green A., Hamp-Lyons, L., Sifakis, N. and Kordia, S. (2018). *Handbook of Assessment for Language Teachers*. Retrieved from <http://taleproject.eu/>. ISBN 978-9925-7399-1-2 (digital).
- Vogt, K., & Tsagari, D. (2014) Assessment literacy of foreign language teachers: Findings of a European study, *Language Assessment Quarterly*, 11(4), 374–402. <https://doi.org/10.1080/15434303.2014.960046>
- Vogt, K., Tsagari, D., & Csépes, I. (2020). Linking learners' perspectives on language assessment practices to teachers' assessment literacy enhancement (TALE): Insights from four European countries. *Language Assessment Quarterly*, 17(4), 410–433. <https://doi.org/10.1080/15434303.2020.1776714>
- 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>
- Wall, D., & Taylor, C. (2014). Communicative language testing (CLT): Reflections on the “Issues Revisited” from the perspective of an examinations board. *Language Assessment Quarterly*, 11(2), 170–185. <https://doi.org/10>

.1080/15434303.2014.902058

- Winterbottom, M., Brindley, S., Taber, K. S., Fisher, L. G., Finney, J., & Riga, F. (2008). Conceptions of assessment: trainee teachers' practice and values. *The Curriculum Journal*, 19(3), 193-213. <https://doi.org/10.1080/09585170802357504>
- Xu, Y., & Liu, Y. (2009). Teacher assessment knowledge and practice: A narrative inquiry of a Chinese college EFL teacher's experience. *TESOL Quarterly*, 43(3), 493-513. <https://doi.org/10.1002/j.1545-7249.2009.tb00246.x>
- Yan, X., Zhang, C., & Fan, J. J. (2018). Assessment knowledge is important, but ...: How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System*, 74, 158-168. <https://doi.org/10.1016/j.system.2018.03.003>



## Appendix 1

### Teacher Questionnaire

#### Part I. General Information

1. Do you work in (country)?	Yes <input type="checkbox"/> No <input type="checkbox"/>
2. What subject(s) do you teach?	
3. What subjects have you studied?	
4. What is your highest qualification?	Please specify:
5. Type of school/institution you teach at:	
6. Average age of pupils:	
7. Your functions at school/institution:	<input type="checkbox"/> Teacher <input type="checkbox"/> Head of department or school <input type="checkbox"/> Mentor <input type="checkbox"/> Advisory function for authorities (local government, ministry etc.) <input type="checkbox"/> Other? Please specify:
8. During your pre-service or in-service teacher training, have you learned something about testing and assessment (theory and practice)?	<input type="checkbox"/> Yes (please specify:) _____ <input type="checkbox"/> No

#### Part II. Questions About Training in LTA

##### 1. Classroom-focused LTA

1.1. Please specify if you were trained in the following domains

	Not at all	A little (1-2 days)	More advanced
Preparing classroom tests			
Using ready-made tests from textbook packages or from other sources			
Giving feedback to students based on information from tests/assessment			
Using self- or peer-assessment			
Using informal, continuous, non-test types of assessment			
Using the European Language Portfolio, an adaptation of it or some other portfolio			

1.2 Please specify if you need training in the following domains

	None	Yes, basic Training	Yes, more advanced training
Preparing classroom tests			
Using ready-made tests from textbook packages or from other sources			
Giving feedback to students based on information from tests/assessment			
Using self- or peer-assessment			
Using informal, continuous, non-test types of assessment			
Using the European Language Portfolio, an adaptation of it or some other portfolio			

**2. Purposes of testing**

2.1. Please specify if you were trained in the following domains

	Not at all	A little (1-2 days)	More advanced
Giving grades			
Finding out what needs to be taught/ learned			
Placing students into courses, programs, etc.			
Awarding final certificates (from school/program; local, regional or national levels)			

2.2. Please specify if you need training in the following domains

	None	Yes, basic Training	Yes, more advanced training
Giving grades			
Finding out what needs to be taught/ learned			
Placing students into courses, programs, etc.			
Awarding final certificates (from school/program; local, regional or national levels)			

**3. Content and concepts of LTA**

3.1. Please specify if you were trained in the following domains

	Not at all	A little (1-2 days)	More advanced
1. Testing/Assessing:			
a) Receptive skills (reading/listening)			
b) Productive skills (speaking/ writing)			
c) Microlinguistic aspects (grammar/vocabulary)			
d) Integrated language skills			
e) Aspects of culture			
2. Establishing the reliability of tests/assessment			
3. Establishing the validity of tests/assessment			
4. Using statistics to study the quality of tests/assessment			

3.2. Please specify if you need training in the following domains

	None	Yes, basic Training	Yes, more advanced training
1. Testing/Assessing:			
a) Receptive skills (reading/listening)			
b) Productive skills (speaking/ writing)			
c) Microlinguistic aspects (grammar/vocabulary)			
d) Integrated language skills			
e) Aspects of culture			
2. Establishing the reliability of tests/assessment			
3. Establishing the validity of tests/assessment			
4. Using statistics to study the quality of tests/assessment			

## Appendix 2

## A Sample Exam

Name: \_\_\_\_\_

Number: \_\_\_\_\_

Class: \_\_\_\_\_

**— Answer Key —**

**B**

**2008-2009 FIRST TERM FIRST WRITTEN EXAM**

When we are in space, we eat, sleep, wash, work and play just like on the earth. Every day is different, everyday is interesting and none of the days are typical.

I usually get up at six o'clock in the morning. I have breakfast with my friends. After breakfast, I do my regular exercise. Doing exercise is very important because in space we can't use our legs and it is very dangerous for us.

Then, I work on my research about space. I do some experiments. After we finish our work, we have lunch. We eat our meals very carefully because food can also float away.

In the evening, I sometimes watch films or play games with my friends in my free time. I always write e-mails to my friends and talk with my family on the radio.

At night, I usually sleep in a sleeping bag. Some of my friends like floating on air, but I hate it. I tie my sleeping bag to a wall so that I don't float during the night.

This is sometimes difficult but mostly enjoyable life. It is interesting to watch the world from the space.

A) Read the passage again and answer the questions. (5x2=10 pts)

- What does she do after breakfast?  
After breakfast, she does her regular exercise.
- Why does she do exercise?  
Because in space, they can't use their legs.
- What does she do after doing exercise?  
She works on her research about space after doing exercise.
- What does she do in her free time?  
She watches films or play games with her friends in her free time.
- How does she communicate with her family?  
She talks with her family on the radio.

B) Answer the questions according to the information chart. (8x2=16 pts)

- Where does he live?  
He lives in Manchester
- What does he do?  
He is a musician
- How many children has he got?  
He has got two children
- Which languages does he speak?  
He speaks English and French
- Does he smoke?  
He doesn't smoke
- Which day doesn't he work?  
He doesn't work on Sunday
- What does he like doing?  
He likes swimming and music
- What does he hate doing?  
He hates football and walking

Name	: Jack Morgan
Hometown	: Manchester
Work	: Musician
Marital status	: married
Children	: 2 girls
Job detail	: - play guitar - sing songs
Language	: English, French
Smoking	: X
Driving licence	: ✓
Free days	: Sunday
Likes	: music, swimming
Dislikes	: football, walking