

Exploring Language Assessment Literacy of EFL Instructors in Language Preparatory Programs

Tuğba KAYA¹

Enisa MEDE²

Abstract

The overall purpose of this study is to investigate language assessment knowledge of English as a Foreign Language (EFL) instructors working in language preparatory programs of the universities in Turkey. The study also aims to find out EFL instructors' perceptions of language assessment and its reflections in their teaching practices. In order to measure their level of language assessment knowledge, Language Assessment Knowledge Scale (LAKS) was administered to 195 EFL instructors from different universities. With regard to the qualitative phase, interviews were conducted with 17 EFL instructors to find out their perceptions of language assessment and its reflections in the classroom. The findings demonstrated that EFL instructors had considerable knowledge of language assessment, and they could reflect their knowledge in their teaching practice. Also, it was found that certain demographic features, including educational background, teaching experience, or attending any professional development training on language assessment, did not have any impact on EFL instructors' language assessment knowledge. The findings offer several implications for EFL instructors regarding language assessment and provide recommendations for further research in the field of testing and assessment.

Keywords: *Testing and assessment; Assessment literacy; Assessment knowledge; EFL*

¹Tuğba KAYA, English Instructor, Istanbul Kadir Has University, tugbakaya05@gmail.com
<https://orcid.org/0000-0001-7534-4965>

²Assoc. Prof. Enisa MEDE, Bahcesehir University, enisa.mede@es.bau.edu.tr, <https://orcid.org/0000-0002-6555-5248>

Makale geliş tarihi / received: 25.06.2020

Makale kabul tarihi / accepted: 02.03.2021

DOI: 10.17932/IAU.EFD.2015.013/efd_v07i008

İngilizce Hazırlık Programlarında Görev Alan Okutmanların Dil Ölçme ve Değerlendirme Bilgilerinin Araştırılması

Öz

Bu çalışmanın temel amacı, Türkiye’deki İngilizce hazırlık programlarında görev yapmakta olan okutmanların dildeki ölçme ve değerlendirme bilgilerinin incelenmesidir. Bu çalışma, aynı zamanda okutmanların dildeki ölçme ve değerlendirmeye yönelik algılarını ve bunun öğretmenlik uygulamalarına yansımalarını bulmayı hedeflemektedir. Dildeki ölçme ve değerlendirme seviyelerini bulmak için, Dilde Ölçme Değerlendirme Bilgisi Ölçeği farklı üniversitelerden 195 İngilizce okutmanına uygulanmıştır. Çalışmanın nitel veri aşamasında ise, dil ölçme ve değerlendirme algıları ve bunun sınıflarına yansımalarını bulmak amacıyla 17 İngilizce okutman ile röportajlar gerçekleştirilmiştir. Sonuçlar, okutmanların yeterli düzeyde ölçme ve değerlendirme bilgilerine sahip olduklarını ve bunu sınıf içi uygulamalarına yansıtılabildiklerini göstermiştir. Ayrıca, eğitim durumu, öğretmenlik deneyimi, dilde ölçme ve değerlendirme konusunda eğitimlere katılıp katılmama gibi demografik özelliklerin katılımcıların ölçme ve değerlendirme bilgilerine bir etkisi olmadığı tespit edilmiştir. Bu bulgular, İngilizce okutmanları için dilde ölçme ve değerlendirme hakkında tavsiyeler ve ölçme ve değerlendirme alanında ileriki araştırmalar için öneriler sunmaktadır.

Anahtar Kelimeler: *Dilde Ölçme ve Değerlendirme, Dilde Ölçme ve Değerlendirme Okuryazarlığı, Dilde Ölçme ve Değerlendirme Bilgisi*

INTRODUCTION

Assessment has a crucial role in education since it allows to improve learning process (Cowie & Bell, 1999). It is also an indispensable part of teaching and learning since it enables language teachers to identify the language level of their students (Taras, 2005). It plays a key role in language education since it is used to develop language instruction and support language learning (Davison & Leung, 2009). Also, teachers can find out the strengths and weakness of their students and act accordingly through the assessment practices (Wojtczak, 2002). Furthermore, assessment literacy allows language teachers to choose the appropriate teaching and assessment strategies which will benefit their students to the utmost (Tosuncuoglu, 2018).

Language Assessment Literacy

The term “assessment literacy” was first coined by Richard Stiggins (1991). According to Stiggins (1991), assessment literacy can be defined as the ability to distinguish poor-quality and high-quality assessments and use it for drawing conclusions about student achievement. Language assessment literacy (LAL) is originated from the general term “assessment literacy”. It can be defined as language teachers’ knowledge of testing and how they implement this knowledge in their classrooms in order to measure language ability (Malone, 2013). Teachers who have a considerable level of assessment literacy are able to make inferences about assessment results, understand students better, and set learning objectives in the long run (Gottheiner & Siegel, 2012). Therefore, having assessment literacy is of great importance for teachers and educators. It is considered to be an indispensable part of educational programs for both teachers and other stakeholders (Taylor, 2009).

In the light of all the facts mentioned above, it is clear that language teachers need to have sufficient amount of language assessment literacy in order to conduct assessment activities in classroom. Nevertheless, there is a gap in the literature as to what language teachers need to know regarding assessment and what they actually know. Most of the studies in the assessment literature indicate that language teachers do not have adequate training in assessment, and they do not have a sufficient level of LAL (Inbar-Lourie, 2017).

LITERATURE REVIEW

Several studies investigating assessment literacy of teachers have been carried out in order to shed some light on the issue (Plake, 1993; Mertler, 1998). Recently, there has been an increasing amount of literature on language assessment literacy of teachers. A brief literature of review on language assessment literacy has focused on four main concepts: teachers’ level of language assessment literacy, assessment-related needs and perceptions of language assessment, and their classroom assessment practices. Language assessment literacy (LAL) has also gained attention in the Turkish context in recent years. Few studies have been published on language assessment literacy of Turkish EFL instructors, and more recent evidence (Öz & Atay, 2017; Mede & Atay, 2017; Öztürk & Aydın, 2018) suggests that Turkish EFL instructors’ language assessment literacy is quite limited.

To begin with, Plake (1993) conducted a national study with 555 teachers from 45 states in the United States so as to measure their assessment literacy, and the results of his study showed that teachers did not have adequate knowledge of assessment practices to measure students' performances. Similarly, Mertler (1998) aimed to investigate classroom assessment practices of teachers in Ohio and gain an insight into different forms of assessment techniques teachers use in their classroom. A survey which was designed by the researcher in order to identify the traditional and alternative assessment techniques teachers use in their classes was administered to 625 teachers working with different levels. The results indicated that teachers do not usually conduct statistical analyses for their assessment data regardless of their teaching experiences.

Some studies have also focused on language assessment knowledge and training needs of EFL teachers. The study conducted by Firoozi et al. (2019) aimed to investigate Iranian EFL instructors' assessment literacy and the relationship between their classroom assessment needs and the assessment policies that are currently implemented in Iran. According to the findings of the study, it was reported that Iranian EFL instructors had limited knowledge of language assessment.

Muñoz et al. (2012) examined EFL instructors' perceptions and practices regarding oral and written assessment through surveys, interviews, and a written report of experiences. They found that there was a mismatch between what teachers believe and what they do regarding language assessment. Similarly, Jannati (2015) also aimed to investigate Iranian ELT teachers' perceptions towards language assessment and their in-class assessment practices based on their teaching experiences. He found that although the instructors are familiar with basic concepts in testing and assessment, they do not reflect their language assessment literacy in their teaching practices. Sultana (2019) also attempted to find out English teachers' perceptions of assessment practices and its reflections in their classroom. Semi-structured interviews were conducted with 10 English teachers from five different schools in Bangladesh. The results showed that none of the participants had adequate amount of training in testing and assessment despite the fact that they are in charge of developing tests to assess students' learning.

Considering the Turkish context regarding assessment literacy, few studies to date have been conducted. Karaman and Şahin (2014) conducted a study in order to find out the perceptions of pre-service English teachers

regarding assessment and their assessment literacy. Their study proved that more training in assessment is needed for pre-service teachers to develop their assessment literacy. Similarly, Hatipoğlu (2015) revealed that although students are trained for four years in the ELT department and take a great number of tests, they still have little knowledge of testing and assessment. Another study conducted by Öz and Atay (2017) investigated Turkish EFL instructors' perceptions towards language assessment and how they reflect their assessment knowledge in their teaching practices. The results showed that although EFL instructors had a certain amount of knowledge of language assessment, they have difficulty in reflecting their knowledge in their classes. Mede and Atay (2017) also conducted a study on 350 EFL instructors working at English preparatory schools in universities with an aim to find out their language assessment knowledge and training needs regarding language assessment and testing. An online questionnaire adapted from Vogt and Tsagari's (2014) study and focus group interviews were utilized to get a better understanding of EFL instructors' perceptions towards language assessment practices in their workplace. The findings revealed that Turkish EFL instructors had limited knowledge of language assessment and testing, and therefore they needed further training in developing tests and providing feedback. Finally, Öztürk and Aydın (2018) conducted a study with 542 EFL instructors in order to investigate general and skill-based language assessment literacy of EFL instructors through the LAKS scale developed by themselves. The results showed that Turkish EFL instructors' language assessment literacy is quite limited.

To conclude, there are a number of studies conducted on language assessment literacy of teachers. However, the number of the studies which attempt to find out language assessment knowledge of EFL instructors is still quite limited in the Turkish context.

The Current Study

The overall purpose of this study is to investigate language assessment knowledge of EFL instructors working at preparatory schools of universities. Particularly, the study has the following sub-objectives:

- to develop a better understanding of EFL instructors' language assessment knowledge level regarding four major language skills
- to determine whether there is a relationship between their language assessment knowledge and certain demographic features they have including educational background, years of experience, subject of

graduation, having a testing and assessment course in undergraduate program and receiving training in testing.

- to find out EFL instructors' perceptions of language assessment literacy and its reflections in their teaching practices.

To meet the objectives of this study, the following research questions were addressed:

1. What is the general and skill-based level of language assessment knowledge (LAK) of EFL instructors working at preparatory schools in Turkey?
2. Is there any relationship between their LAK level and the following demographic features?
 - 2a. years of teaching experience
 - 2b. educational background
 - 2c. subject of graduation
 - 2d. having a separate testing and assessment course in pre-service education
 - 2e. attending any professional development training on language assessment
 - 2f. being a testing office member or not
3. What are the perceptions of EFL instructors about their classroom-based language assessment practices?
4. How do they reflect their LAK in their teaching practices?

METHOD

Research Design

A mixed-method approach was adopted in this study by combining qualitative and quantitative methods in order to examine EFL instructors' both level of LAK and their perceptions towards language assessment practices. Quantitative data were collected via LAKS in order to investigate EFL instructors' level of general and skill-based language assessment knowledge. The quantitative data, on the other hand, were collected from interviews conducted with EFL instructors working in higher education

context in Turkey.

Participants

The participants of the study were EFL instructors working at the preparatory schools of universities in Istanbul, Turkey. The participants who took part in this study were graduates of several departments which are related to English language. In total, 195 instructors from 37 universities in Istanbul participated in the study. The following table shows the number of the participants and their demographic features.

Table 1. Distribution of the participants based on their certain demographic features

Years of experience	n	%
1-5 years	58	29.7
6-10 years	61	31.3
11-15 years	28	14.4
16-20 years	20	10.3
More than 21 years	28	14.4
Total	195	100
Subject of graduation from		
ELT	101	51.8
Non-ELT	94	48.2
Total	195	100
Educational Background		
BA degree	72	36.9
MA degree	111	56.9
PhD degree	12	6.2
Total	195	100
Being a testing office member		
Yes	68	34.9
No	127	65.1
Total	195	100
Having a separate testing and assessment course in pre-service education		
Yes	114	58.5
No	81	41.5
Total	195	100

Attending any professional development training on language assessment		
Yes	136	69.7
No	59	30.3
Total	195	100

The qualitative part of the study, on the other hand, included interviews with 17 EFL instructors working at the preparatory school of a foundation university. All the instructors willing to take part in the study were graduates of different BA programs including English Language Teaching, English Language and Literature, American Culture and Literature, and Translation Studies. Their teaching experiences ranged between 1-15 years.

Data collection instruments

Since the study involved a mixed-method approach, both quantitative and qualitative data collection instruments were used in the study. For the quantitative part of the study, a questionnaire named LAKS which was designed to measure LAK of EFL instructors was used. The qualitative part of the study included interviews with EFL instructors regarding their perceptions of language assessment practices.

Language Assessment Knowledge Scale (LAKS)

In order to measure language assessment knowledge of EFL instructors, LAKS developed by Öztürk and Aydın (2018) was adopted and administered to the participants. The questionnaire consists of two parts in total. The first part is the demographic information part, in which participants are required to provide necessary information regarding their demographic features. The second and the main part of the questionnaire consists of 60 items, which were designed to measure participants' assessment knowledge of four language skills. For each correct answer the participants choose, they would get "1", and for each incorrect answer or the option "don't know", they would get "0". Therefore, the highest score that can be achieved for each language skill is 15, whereas the highest score that can be achieved for the total is 60.

Open-ended interview questions

Since the aim of the qualitative part of the study was to explore EFL instructors' perceptions of language assessment and its reflections in the classroom, interviews were conducted with EFL instructors working at the preparatory school of a foundation university in Istanbul, Turkey. The

interview questions that were used in the study were adapted from Jannati (2015). In total, the interview included 10 open-ended questions that were designed to collect data to get an insight into teachers' perceptions of language assessment and their in-class practices in respect of language assessment.

Data Analysis Procedures

In the quantitative part of the study, the data obtained from the questionnaires were analysed through SPSS. Participants' level of general and skill-based language assessment knowledge was analysed through descriptive statistics, whereas the relationship between their certain demographic features and their LAK level was analysed through inferential statistics. Primarily, independent samples t-test was used in order to figure out the relationship between participants' LAK level and their certain demographic features including the BA programme they graduated, having a separate testing course in pre-service education or not and attending trainings on testing and assessment or not. In addition, Pearson Correlation was used in order to examine the relationship between participants' LAK level for each skill. As for the qualitative part, the data were analysed by means of content analysis.

Reliability and Validity

The following table presents the Cronbach Alpha coefficients of LAKS.

Table 2. Reliability Analysis for Language Assessment Knowledge Scale (LAKS)

Reliability Statistics	
Cronbach's Alpha	N of Items
0.83	60

As indicated in Table 2, Cronbach Alpha coefficient of LAKS was found to be 0.83, which shows that it has a statistically high reliability to be used as a measurement tool.

In order to ensure validity of the qualitative findings, peer debriefing and member checking methods were used in the qualitative part of the study. At the end of qualitative data collection period, the findings were shared with the participants so as to make sure that the interpretations made by the researcher did not include any inaccuracies. As for debriefing, the raw data was shared with colleague and discussions were held to reach a consensus on the accuracy of the researcher's interpretations or explanations.

RESULTS

Quantitative Data Analysis

The following section presents participants' general and skill-based language assessment knowledge derived from their responses.

Table 3. General and skill-based LAK level of EFL instructors working at the preparatory programmes of universities in Istanbul, Turkey

N=195	M	SD
Reading	13.93	2.132
Listening	12.05	1.675
Writing	11.16	1.675
Speaking	8.86	1.370
TOTAL	46.00	6.852

Table 3 displays that participants' mean score out of 60 items is 46,00, which means that out of 60 items, the participants answered 46 of the items correctly on average. For the purpose of finding out whether this score is statistically and significantly higher than the half score, one sample t-test was used. The lowest score that can be obtained from the scale is 0, whereas the highest score that can be achieved is 60. Therefore, 30, which is half of the total score, was accepted as the reference point.

Table 4. One sample t-test results

M	MD	df	t	p
46,00	15,00	194	31,674	0.000*

* $p < 0,05$ Test Value = 30

Table 4 shows that the mean difference between participants' mean score ($M = 46.00$) and half of the score which was accepted reference point (30) is 15.00, which means that it is statistically significant. These findings show that their language assessment knowledge in general is significantly high.

The first question in the study also aimed to investigate skill-based LAK level of EFL instructors working at the preparatory programmes in universities in Istanbul. The highest score that can be achieved for each skill is 15, and the lowest score is 0. Hence, half of the total score was 7.5. The participants who obtained a higher mean score than 7.5 for each skill was regarded as knowledgeable about the related skill.

Table 5. Reading LAK level of EFL instructors working at the preparatory programmes in Istanbul

ASSESSING READING	T (%)	F (%)	DKW (%)	M	SD
1. Asking learners to summarize the reading text is a way of assessing their reading skills.	85.60	10.80	3.60	1.18	0.469
2. When asking several questions about a reading text, all the questions are independent of each other.	44.60	47.70	7.70	1.63	0.623
3. Cloze test is used for assessing the main idea of the text.	16.90	70.30	12.80	1.96	0.545
4. In a reading exam, using a text learners have encountered before is not a problem.	22.10	66.20	11.80	1.90	0.574
5. One reading text is enough to be included in a reading exam.	22.10	67.70	10.30	1.88	0.557
6. The language of the questions is simpler than the text itself.	68.70	20.00	11.30	1.43	0.687
7. Errors of spelling are penalized while scoring.	21.50	69.20	9.20	1.88	0.542
8. Taking vocabulary difficulty into consideration is necessary in assessing reading skills.	87.20	7.20	5.60	1.18	0.515
9. Including not stated/doesn't say along with true/false items has advantages over true/false items.	69.70	13.30	16.90	1.47	0.769
10. The more items a reading text is followed, the more reliable it becomes.	41.50	31.80	26.70	1.85	0.814
11. Using the same words in the correct option as in the text is not a problem.	18.50	70.30	11.30	1.93	0.542
12. Simplification of reading texts is avoided.	25.10	60.00	14.90	1.90	0.626
13. Reading texts in a reading exam include various genres (essay, article, etc.)	88.20	6.20	5.60	1.17	0.508
14. In top-down approach, assessment is on overall comprehension of the reading text.	73.30	7.70	19.00	1.46	0.794
15. Using ungrammatical distractors in multiple-choice questions in a reading exam is a problem.	80.00	13.30	6.70	1.27	0.575
READING TOTAL (N=195)				13.93	1.067

Note: T: True; F: False; DKW: Don't Know

Table 5 displays that the mean score of the participants' LAK level for reading skill is 13.93, which shows that it is significantly higher than the half score (7.5). These findings show that the participants are quite competent in assessing reading skill.

The participants' LAK level for listening skill is presented in Table 6. The bold numbers refer to the number of the participants who answered the item correctly.

Table 6. Listening LAK level of EFL instructors working at the preparatory programmes in Istanbul

ASSESSING LISTENING	T (%)	F (%)	D K W M (%)	SD	
16. Using reading texts for listening purposes poses a problem.	55.40	33.80	10.80	1.55	0.682
17. Including redundancy (e.g. what I mean to say is that....) in a listening text poses a problem.	19.50	74.90	5.60	1.86	0.483
18. Any type of listening text is used for note-taking.	15.40	77.90	6.70	1.91	0.463
19. Spelling errors are ignored in scoring the dictation.	32.80	58.50	8.70	1.76	0.599
20. Errors of grammar or spelling are penalized while scoring.	34.90	55.90	9.20	1.74	0.614
21. A listening cloze test is a way of selective listening.	73.80	5.60	20.50	1.47	0.814
22. Phonemic discrimination tasks (e.g. minimal pairs such as sheep-ship) are examples of integrative testing.	35.40	18.50	46.20	2.11	0.899
23. Scoring in note-taking is straightforward.	32.30	42.10	25.60	1.93	0.760
24. In discrete-point testing, comprehension is at the literal/local level.	32.80	10.80	56.40	2.24	0.917
25. Using dictation diagnostically in assessing listening skills does not pose a problem.	32.80	36.40	30.80	1.98	0.799
26. Giving learners a transcript of the listening text is a valid way of assessing listening skills.	9.70	83.10	7.20	1.97	0.412
27. Dictation is a kind of discrete-point testing.	34.40	14.90	50.80	2.16	0.910
28. Inference questions based on intelligence are avoided in listening tests.	68.70	21.00	10.30	1.42	0.671
29. Asking learners to listen to names or numbers is called intensive listening.	34.90	39.50	25.60	1.91	0.774

30. In selective listening, learners are expected to look for certain information.	92.30	1.00	6.70	1.14	0.508
LISTENING TOTAL (N= 195)				12.05	2.949

Note: T: True; F: False; DKW: Don't Know

Table 6 shows that the mean score of the participants' LAK level for listening skill is 12.05, which is significantly higher than the half score. Therefore, it can be said that the EFL instructors working at the preparatory schools of the universities in Istanbul have considerable knowledge of assessing listening skill.

The following table displays the LAK level of the participants for writing skill. The bold numbers refer to the number of the participants who answered the item correctly.

Table 7. Writing LAK level of EFL instructors working at the preparatory programmes in Istanbul

ASSESSING WRITING	T (%)	F (%)	DKW (%)	M	SD
31. Giving two options to learners and asking them to write about one ensure reliable and valid scoring.	63.60	23.10	13.30	1.50	0.721
32. Analytic scoring is used to see the strengths and weaknesses of learners.	76.90	7.70	15.40	1.38	0.740
33. The parts of a scoring scale and the scores in each part do not change for different levels of learners.	26.70	62.60	10.80	1.84	0.592
34. When there is a disagreement between the scores of the two raters, they score the written work again."	69.20	26.20	4.60	1.35	0.568
35. Learners are required to write about at least two tasks in the exam rather than one task.	28.20	56.40	15.40	1.87	0.649
36. Giving restrictive prompts/guidelines to learners for the writing task is avoided.	35.40	56.90	7.70	1.72	0.597

37. Giving learners an opinion and asking them to discuss it is a valid way of assessing their writing skills.	73.80	17.90	8.20	1.34	0.626
38. Using visuals which guide learners for writing poses a problem.	14.90	68.70	16.40	2.02	0.561
39. Holistic scoring is used to see whether the learner is proficient or not at the end of the term.	61.50	16.90	21.50	1.60	0.821
40. Analytic scoring leads to greater reliability than holistic scoring in writing.	55.90	13.80	30.30	1.74	0.894
41. In controlled writing, learners have the chance to convey new information.	28.70	47.70	23.60	1.95	0.723
42. Classroom evaluation of learning in terms of writing is best served through analytic scoring rather than holistic scoring.	51.80	17.40	30.80	1.79	0.886
43. Irrelevant ideas are ignored in the assessment of initial stages of a written work in process writing.	35.90	53.80	10.30	1.74	0.631
44. Providing a reading text for writing is a way of assessing writing skills.	54.40	29.20	16.40	1.62	0.753
45. Mechanical errors (e.g. spelling and punctuation) are dealt with in the assessment of later stages of a written work.	49.70	43.10	7.20	1.57	0.625
WRITING TOTAL (N= 195)				11.16	3.838

Note: T: True; F: False; DKW: Don't Know

According to Table 7, the mean score of the participants in assessing writing is 11.6. Therefore, it can be stated that their LAK level for writing skill is statistically high.

Finally, the LAK level of the participants for assessing speaking skill is presented in the following table.

Table 8. Writing LAK level of EFL instructors working at the preparatory programmes in Istanbul

ASSESSING SPEAKING	T (%)	F (%)	D K W M SD (%)		
46. When the interlocutor does not understand the learner, giving that feeling or saying it poses a problem.	57.90	31.80	10.30	1.52	0.676
47. Giving learners one task is enough to assess speaking skills.	9.70	86.20	4.10	1.94	0.369
48. Interlocutors' showing interest by verbal and non-verbal signals poses a problem.	44.60	48.20	7.20	1.63	0.616
49. When it becomes apparent that the learner cannot reach the criterion level, the task is ended.	40.50	49.20	10.30	1.70	0.647
50. Using holistic and analytic scales at the same time poses a problem.	31.30	37.40	31.30	2.00	0.793
51. Reading aloud is a technique used to assess speaking skills.	25.10	64.60	10.30	1.85	0.577
52. In interlocutor-learner interviews, the teacher has the chance to adapt the questions being asked.	64.10	25.60	10.30	1.46	0.675
53. In interactive tasks, more than two learners pose a problem.	31.30	55.90	12.80	1.82	0.640
54. The interlocutor gives the score when the learner is in the exam room.	11.30	81.00	7.70	1.96	0.435
55. In a speaking exam, production and comprehension are assessed together.	72.30	20.50	7.20	1.35	0.610
56. Asking learners to repeat a word, phrase or a sentence is a way of assessing speaking skills.	30.30	56.40	13.30	1.83	0.640
57. Discussion among learners is a way of assessing speaking skills.	83.60	10.30	6.20	1.23	0.547
58. A checklist is a means of scoring oral presentations in in-class assessment.	82.60	5.10	12.30	1.30	0.676

59. When the focus is to assess discourse, role-plays are used.	67.20	11.30	21.50	1.54	0.826
60. In peer interaction, random matching is avoided.	21.00	54.90	24.10	2.03	0.673
SPEAKING TOTAL (N= 195)				8.86	6.138

Note: T: True; F: False; DKW: Don't Know

As shown in Table 8, the participants had a mean score of 8.86 in assessing speaking, which is higher than the half score. Thus, their LAK level for speaking skill is regarded to be significantly high.

The second question in the study aimed to explore whether there was a relationship between the participants' LAK level and certain demographic features.

The first variable to be investigated was years of teaching experience. One-way ANOVA was used in order to see if there is any relationship between the participants' teaching experience and their LAK level.

Table 10. Participants' Language Assessment Knowledge Based on Teaching Experience

	Sum of Squares	df	Mean Square	F	p
Between Groups	45.131	4	11.283	0.233	0.920
Within Groups	9200.85	190	48.426		
TOTAL	9245.981	194			

Table 10 shows that the value of F is 0.233, which does not reach a significance with a p-value of 0.920 ($p = 0.92 > .05$). Therefore, it was concluded that there was not a statistically significance between groups in terms of their teaching experience.

The second variable to be investigated was educational background of the participants. In order to investigate whether a relationship exists or not between the participants' LAK level and having a BA degree, MA degree or PhD degree, one-way ANOVA was utilized.

Table 11. Participants' Language Assessment Knowledge Based on Educational Background

	Sum of Squares	df	Mean Square	F	p
Between Groups	47.377	2	23.689	0.494	0.611
Within Groups	9198.604	192	47.909		
TOTAL	9245.981	194			

Table 11 shows that the F value (0.464) does not reach a significance with a p value of 0.611 ($p = 0.611 > .05$). These findings suggest that there was no relationship between the participants' educational background and their LAK level.

The third variable to be investigated in the study was the BA programme the participants graduated from. Independent samples t-test was used in order to find out whether there is a relationship between the participants' LAK level and having graduated from an ELT department or a non-ELT department. The table below shows the frequency of the participants based on the BA programme they graduated and the mean scores of their LAK level.

Table 12. Participants' Language Assessment Knowledge Based on Subject of Graduation

Subject of Graduation	N	M	t	df	p
ELT	101	45.00	-1.374	103	0.171
non-ELT	94	46.36			

Table 12 shows that out of 195 participants, 101 of them were graduates of ELT department while 94 of the participants were graduates of non-ELT departments. The p value (0.171) showed that there was no statistically significant difference between those who graduated from ELT department and those who graduated from a non-ELT department ($p = 0.171 > .05$). Therefore, it can be interpreted that the BA programme the participants graduated from did not have an impact on their LAK level.

Having a testing and assessment course in undergraduate education was the fourth variable to be investigated. In order to examine the relationship between the participants' LAK level and having received a testing and assessment course in pre-service education, independent samples t-test was utilized. The findings are presented in the following table.

Table 13. Participants' Language Assessment Knowledge Based on Having a Separate Testing and Assessment Course in Pre-service Education

Having a separate testing course	N	M	t	df	p
Yes	114	44.8076	-2.06	193	0.041*
No	81	46.8574			

* $p < 0,05$

In table 13, the p value (0.041) shows that there was a statistically significant difference between those who had a testing course and those who did not in their undergraduate studies ($p = 0.041 < .05$). According to these findings, the LAK level of those who had a testing and assessment course in their pre-service education was found to be slightly lower than those who did not have any testing course in their pre-service education. However, the mean difference between the two groups was quite low (2.05).

The fifth variable that was investigated in the study was having trainings on testing and assessment. In order to find out whether there was any relationship between those who received trainings and those who did not, independent samples t-test was used. The findings are shown in the following table.

Table 14. Participants' Language Assessment Knowledge Based on Attending any Professional Development Training on Language Assessment

Attending any professional development training on language assessment	N	M	t	df	p
Yes	136	45.5	-0.481	193	0.631
No	59	46.02			

The p value (0.631) showed that no statistically significant difference was found between the two groups ($p = 0.631 > .05$). It means that receiving a testing and assessment training did not have an impact on the participants' LAK level.

Finally, in order to see whether any relationship exists between the LAK level of the participants who have been a testing office member and the LAK level of those who have not, independent samples t-test was utilized. The findings are shown below in Table 15.

Table 15. Participants' Language Assessment Knowledge Based on Being a Testing Office Member or Not

Being a testing office					
member or not	N	Mean	t	df	p
Yes	68	46.11	0.67	193	0.504
No	127	45.41			

According to table 15, the p value (0.504) shows that there was no statistically significant difference between the LAK level of the two groups ($p = 0.504 > .05$). That is to say, being a testing office member or not did not have any impact on the LAK level of the participants.

In short, the findings suggest that there was no statistically significant relationship between the participants' LAK level and the following variables: teaching experience, educational background, subject of graduation, having trainings on testing and assessment, and being a testing office member or not. Among all the variables, the only variable which had an impact on the participants' LAK level was having a testing and assessment course in undergraduate education. Interestingly, those who had no testing course in their pre-service education had a higher level of LAK. However, despite the difference, their LAK levels were still very close to each other.

Qualitative Data Analysis

The third question in the study aimed to investigate the participants' perceptions of their language assessment practices. In order to find out their perceptions, interviews were conducted with EFL instructors. When the participants were asked about their opinions on the characteristics of a good assessment, they listed a number of features to define characteristics of a good assessment in their terms. Table 18 presents the features of a good assessment from the perspectives of the participants and the frequency of all the answers.

Table 16. Characteristics of a good assessment

Characteristics	N
Practical	2
Objective	2
Flexible	1
Fair	2
Justifiable	1
Valid	8
Authentic	1
Reliable	5
Adaptable	1
Constructive	1
Organized	1

The results in Table 16 indicate that the most commonly given answers provided by the teachers were reliability and validity, which means that most of the participants are familiar with the basic terms in assessment. Also, the qualities mentioned by the respondents prove that all the participants have some opinions regarding the basics of an assessment. Therefore, it can be concluded that the participants have some knowledge of basic assessment principles and needs of a good assessment.

When the participants were asked whether students should be informed of what they will be assessed or not, all the participants stated that students need to be informed of what they will be assessed on by indicating different reasons. Some of the reasons mentioned by the participants are listed below:

“Yes, they should be informed, because we should test individual things at a time and students should be aware of the things they should pay attention to.” (Interview Data, 05.07.2019)

“Yes sure, one of the important qualities of a good exam is to have content validity. If students did not know what they will be assessed on, that exam would not have any validity.”

(Interview Data, 02.07.2019)

The findings also showed that participants make use of both formative and summative assessment methods in their classes. Also, written and oral assessment methods, writing tasks, portfolios and pop-up quizzes were mentioned by the participants as activities that are used in classroom to assess students. Finally,

for the question regarding the participants' engagement in peer assessment and self-assessment, the majority of the participants stated that they make use of peer assessment and self-assessment methods in their classroom in order to detect students' weaknesses and strengths.

In brief, the findings suggested that the participating teachers have knowledge of summative assessment and formative assessment methods. Also, they are well-informed about the importance of alternative forms of assessment such as portfolios, peer-assessment and self-assessment. Also, when it comes to in-class assessment practices, it was seen that they engage in both summative and formative assessment methods in their classes.

DISCUSSION

The current study aimed to find out general and skill-based language assessment knowledge level of EFL instructors working at the preparatory schools of the universities in Istanbul, Turkey. The quantitative findings showed that they had a high level of language assessment knowledge.

When the literature is reviewed, it can be seen that the majority of the existing studies on language assessment found teachers had a low level of language assessment. To begin with, Hatipoğlu (2015) conducted a study on 124 pre-service EFL instructors in an ELT department to explore their knowledge of language assessment. The findings of her study showed that pre-service teachers had little knowledge of testing and assessment. In another study conducted by Mede and Atay (2017), it was found that Turkish EFL instructors had limited knowledge of language assessment and testing. Similarly, Öztürk and Aydın (2018) conducted a study to find out language assessment literacy of EFL instructors through the LAKS scale, which was also used in this study. The findings of their study showed that Turkish EFL instructors' language assessment literacy is quite limited. When the findings of the current study are compared to the findings of the previous studies on language assessment, it can be seen that these findings contradict the findings of the existing studies in literature.

The inconsistency between this study and the previous studies might derive from the fact that the current study included many participants from universities with no testing office. In these universities, since there is no separate testing office, all the instructors working at the institution are given regular professional development trainings on testing and assessment to improve the quality of the exams. Therefore, they may be more knowledgeable about testing and assessment than those who do not attend assessment trainings on a regular basis. This might be a significant reason for the discrepancy between the previous studies and the current study.

When the participants' skill-based LAK level was analysed, it was seen that the participants were found to be competent in assessing four language skills, which

contradicts the findings of Öztürk and Aydın's study (2018). In the current study, EFL instructors were found to be competent in assessing four language skills. Also, in the current study, the participants were found to be most competent in assessing reading, whereas they were least competent in assessing speaking among four language skills. The reason for the low scores in assessing speaking may be that, as Alderson and Bachman (2004) suggest, speaking is the most difficult language skill to assess reliably since there is a range of personal factors affecting the judgement of the interlocutor conducting the speaking exam. Similarly, Alderson, Clapham, and Wall (1995) argue that "the success of a speaking test relies on the individuals who are administering the test doing their job well" (p. 116). When all these findings are taken into account, it can be concluded that assessing speaking requires more attention than assessing other language skills.

The second research question in the study aimed to find out whether there is any relationship between the EFL instructors' LAK level and certain demographic features.

Years of teaching experience was the first variable that was examined in the current study. The findings pointed out that teaching experience had no impact on EFL instructors' language assessment knowledge, which shows parallelism with the findings of the previous studies in the literature (Tao, 2014; Jannati, 2015; Öz and Atay; 2017; Hakim, 2015; Öztürk and Aydın, 2018).

Another variable that was investigated in the study was educational background. The findings suggested that there was no relationship between the participants' educational background and their LAK level. These findings are in line with some of the existing studies in the literature (Tao, 2014; Öztürk and Aydın, 2018).

With regard to subject of graduation, no significant difference was found between the EFL instructors' LAK level and the BA programme they graduated from, which contradicts the findings of Tao's study (2014). He found that English-major instructors had a higher level of assessment knowledge than those who did not major in English in their BA education. However, Öztürk and Aydın (2018) also investigated the effect of subject of graduation on EFL instructors' language assessment literacy, and found that having graduated from an ELT department or a non-ELT department did not have any influence on EFL instructors' LAK level, which conforms to the findings of the current study.

Having a separate testing and assessment course in pre-service education was another variable that was examined in the current study. According to the findings, those did not have a testing course in their BA education had a higher level of LAK than those who had a testing course. According to the previous studies in the literature, no relationship was found between pre-service assessment education and teachers' language assessment knowledge (Tsagari, 2008; Tao, 2014; Öztürk and Aydın, 2018). The results of the present study may derive from several

factors. Most importantly, the assessment related courses offered in pre-service education are generally theory-oriented. In other words, pre-service teachers do not receive training on how to prepare test items. Instead, theoretical aspects of testing and assessment are being taught to pre-service teachers. Therefore, they are not equipped with sufficient knowledge of testing and preparing items when they graduate from university. This reason may have led to the findings of the current study.

The current study also investigated the relationship between attending professional development trainings on language assessment and EFL instructors' LAK level. The results suggested that attending in-service assessment trainings did not influence EFL instructors' LAK level, which was in line with the findings of Öztürk and Aydın's study (2018). Mede and Atay (2017) also focused on assessment training needs of EFL instructors, and they found that majority of the participants did not have sufficient training in testing and assessment. When the findings of the current study are considered, the ineffectiveness of attending trainings on testing in the current study might derive from several factors. One reason may be that in-service assessment trainings that teachers attend could be insufficient, since they are mostly 'one-shot trainings' (Mede and Atay, 2017, p. 58). It is simply not possible to be competent in testing and assessment with a short-term training. Besides, the quality of the trainings offered may be another reason for the inefficiency of those trainings. Cumming (2009) argues that neither teacher preparation programs nor professional development opportunities enable teachers to get prepared to conduct assessment tasks. According to Taylor (2013), in order to tackle with language assessment literacy, it is significant to balance the theoretical aspects and practical aspects of language assessment as well as considering ethical issues. Inbar-Lourie (2008) discusses that little research has been conducted on the content and objectives of language assessment courses so far. Therefore, it can be argued that more research is still needed to reach a particular conclusion.

Finally, the last variable that was examined in the current study was being a testing office member or not. The findings showed that having worked as a testing office member did not affect the participants' level of language assessment knowledge. These findings contradict the findings of Öztürk and Aydın's study (2018).

The third question in the study aimed to examine EFL instructors' perceptions of language assessment practices and its reflections in their classroom. The results of the study revealed that EFL instructors are aware of the importance of assessment and its profound impact on teaching. Moreover, they are informed about summative assessment and formative assessment methods including alternative forms of assessment such as peer-assessment, self-assessment and portfolios. These findings are in line with some of the existing studies in the literature (Jannati, 2015; Öz & Atay, 2017).

The fourth question in the study aimed to explore how EFL instructors reflect their language assessment knowledge in their classes. The findings showed that the participants make use of both formative and summative assessment methods to evaluate students' performance. However, these findings contradict the findings of some of the existing studies in the literature (Muñoz, Palacio, & Escobar, 2012; Jannati, 2015). To begin with, Muñoz et al. (2012) found that teachers do not reflect their language assessment knowledge in their teaching practices. Similarly, Jannati (2015) also found that there was a mismatch between what EFL instructors believe and what they actually implement in their classes. The reason for the inconsistency between the current study and the previous studies could be that assessment for learning has recently gained popularity among language teachers and educators. Thus, the kind of activities teachers use in order to evaluate students may have changed over the last few years. The present study showed that EFL instructors can reflect their language assessment knowledge in their teaching practices. However, more research is still needed to get an in-depth understanding of EFL instructors' in-class assessment practices.

PRACTICAL IMPLICATIONS

The present study aimed to shed light on language assessment literacy of EFL instructors working in higher education context. In the light of the quantitative and qualitative findings of the study, the following implications are suggested.

Firstly, the study found that having a testing and assessment course in undergraduate education did not have any impact on EFL instructors' LAK level, which shows that the testing and assessment courses offered in pre-service education do not improve language assessment literacy of teacher candidates. Therefore, more research into the effectiveness of the assessment courses offered in pre-service education should be conducted, and the possible weaknesses these courses could be examined and improved accordingly.

Secondly, the research has shown that ELT graduates do not perform better than non-ELT graduates with regards to their language assessment knowledge despite having a separate testing and assessment course in their undergraduate education. Hence, more courses on testing and assessment should be offered in ELT departments so that pre-service language teachers can be equipped with sufficient knowledge of language assessment when they have graduated from university.

Thirdly, the findings of the research have indicated that having trainings on testing and assessment did not influence EFL instructors' language assessment knowledge. Hence, more research into the effectiveness of assessment trainings offered to in-service language teachers should be done to improve the quality of these trainings.

In short, the findings of the current study shed light on EFL instructors' language

assessment knowledge and their perceptions of language assessment. The implications suggested above may be of great importance for language teachers to improve the standards of teaching and learning by focusing on teachers' language assessment literacy.

CONCLUSIONS

The overall purpose of the current study was to explore general and skill-based language assessment knowledge of EFL instructors working in higher education context in Istanbul, Turkey. The findings of quantitative data showed that EFL instructors' general and skill-based LAK level was pretty high. Also, the study found that certain demographic features such as educational background, having a testing and assessment course in BA education, or having assessment trainings did not have any impact on the participant teachers' LAK level. Therefore, it was concluded that the quality of the testing and assessment courses offered in pre-service education could be improved.

Finally, based on the qualitative findings of the current study, it was concluded that EFL instructors were aware of the importance of assessment in learning process, and they were familiar with summative and formative assessment methods. The findings also showed that they were able to reflect their language assessment knowledge in their classroom. Therefore, it was concluded that their perceptions of language assessment were in line with their in-class teaching practices.

REFERENCES

- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge, England: Cambridge University Press.
- Alderson, J.C., & Bachman, L. (2004). Series editors' preface. In S. Luoma (Eds.), *Assessing speaking* (pp. 9-11). Cambridge: Cambridge University Press.
- doi:10.1017/CBO9780511733017.001
- Cowie, B., & Bell, B. (1999). A model of formative assessment in Science Education. *Assessment in Education: Principles, Policy & Practice*, 6(1), 101-116.
- Firoozi, T., Razavipour, K. & Ahmadi, A. The language assessment literacy needs of Iranian EFL teachers with a focus on reformed assessment policies. *Lang Test Asia* 9, 2 (2019). <https://doi.org/10.1186/s40468-019-0078-7>.
- Gottheiner D. M., & Siegel M. A. (2012). Experienced Middle School Science Teachers' Assessment Literacy: Investigating Knowledge of Students' Conceptions in Genetics and Ways to Shape Instruction. *The Association for Science Teacher Education, USA J Sci Teacher Educ* 23, 531-557. <http://dx.doi.org/10.1007/s10972-012-9278z>

- Hakim, B. (2015). English language teachers' ideology of ELT assessment literacy. *International Journal of Education & Literacy Studies*, 3(4), 42-48.
- Hatipoğlu, Ç. (2015). English language testing and evaluation (ELTE) training in Turkey: Expectations and needs of pre-service English language teachers. *ELT Research Journal*, 4(2), 111-128.
- Inbar-Lourie, O. (2008a). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385-402.
- Inbar-Lourie, O. (2008b). Language assessment culture. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed.) (pp. 285-299). Springer Science Business Media LLC.
- Inbar-Lourie, O. (2017). Guest editorial to the special issue on language assessment literacy. *Language Testing*, 30(3), 301-307.
- Jannati, S. (2015). ELT Teachers' Language Assessment Literacy: Perceptions and Practices. *The International Journal of Research in Teacher Education*, 6(2), 26-37.
- Karaman, P. & Şahin, Ç. (2017). Öğretmen Adaylarının Ölçme-Değerlendirme Okuryazarlıklarının Mikro-Öğretim Yoluyla Geliştirilmesi. *Journal of Turkish Studies*, 12, 255-274.
- Katz, A. & Gottlieb, M. (2013). Assessment in the Classroom. In the Encyclopedia of Applied Linguistics, C. A. Chapelle (Ed.). doi:10.1002/9781405198431.wbeal0042
- Malone, M. (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), 1-5.
- Mertler, C. A. (1998). Classroom assessment practices of Ohio teachers. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Chicago, IL, October (ERIC Document Reproduction Service No. 428 085).
- Muñoz, A. P., Palacio, M., & Escobar, L. (2012). Teachers' beliefs about assessment in an EFL context in Colombia. *Profile: Issues in Teachers' Professional Development*, 14(1).
- Ölmezer-Öztürk, E. & Aydın, B. (2018). Developing and Validating Language Assessment Knowledge Scale (LAKS) and Exploring the Assessment

- Knowledge of EFL instructors. Unpublished PhD Dissertation, Anadolu University, Turkey.
- Ölmezer-Öztürk, E., & Aydın, B. (2019). Investigating language assessment knowledge of EFL instructors. *Hacettepe University Journal of Education*, 34(3), 602-620. doi: 10.16986/HUJE.2018043465
- Öz, S. & Atay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: perceptions and practices. *ELT Research Journal*, 6(1), 25-44.
- Plake, B. S. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Mid-Western Educational Researcher*, 6(1), 21-27.
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kaplan*, 72(7), 534-539.
- Sultana, N. (2019). Language assessment literacy: an uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*. 9. 10.1186/s40468-019-0077-8.
- Tao, N. (2014). *Development and validation of classroom assessment literacy scales: English as a Foreign Language (EFL) teachers in a Cambodian Higher Education Setting*. Unpublished PhD dissertation, Victoria University, Australia.
- Taras, M. (2005). Assessment -summative and formative- some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466-478.
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29(1), 21-36.
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403-412.
- Tosuncuoğlu, I. (2018). Importance of Assessment in ELT. *Journal of Education and Training Studies*, 6(9), 163-167.
- Tsagari, D. (2008). Assessment literacy of EFL instructors in Greece: Current trends and future prospects. PowerPoint presentation at the 5th Annual EALTA Conference, May 9-11, Athens, Greece.
- 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.
- William, D. (2013). Assessment: The Bridge between Teaching and Learning. *Voices from the Middle*, 21(2), 15-20.
- Wojtczak, A. (2002). Medical education terminology. *Medical Teacher*, 24(4), 357-357.