UNDERSTANDING THE DEVELOPMENT OF LANGUAGE ASSESSMENT LITERACY[1]

Ahmet Erdost YASTIBAŞ[2], Mehmet TAKKAÇ[3]

Abstract

Language assessment literacy (LAL) has become an important competence for English language (EL) teachers to have; however, several studies conducted in the literature indicate that Turkish EL teachers have low levels of LAL. Those studies do not focus on how EL teachers develop their LAL. Therefore, the present study aims to find out how Turkish EL teachers have developed their LAL. The study was designed as a small scale qualitative study. Eight EL teachers working at a Turkish university participated into the study. The researchers used a focus group discussion and semi-structured individual interviews to collect data. The researchers made the study trustworthy through member check, triangulation, and thick description. The data collected were content-analyzed. The findings of the study point out that previous assessment experience, assessment training, and self-improvement influence the development of LAL in the participant EL teachers. Self-improvement in which the participant EL teachers peer-assessed their exams, practiced their theoretical information and gained experience by assessing and evaluating their students is found to be the most effective in developing LAL.

Keywords: Language assessment literacy, development, English language teachers

----

YABANCI DİLDE ÖLÇME DEĞERLENDİRME OKURYAZARLIĞININ GELİŞİMİNİ ANLAMAK

Öz

Yabancı dilde ölçme değerlendirme okuryazarlığı, İngilizce öğretmenlerinin sahip olması gereken önemli bir yeterliktir; ancak alınyazında yapılan çalışmalar şunu göstermiştir ki Türk İngilizce öğretmenleri, düşük düzeyde ölçme değerlendirme okuryazarıdır. Bu çalışmalar İngilizce öğretmenlerinin ölçme değerlendirme okuryazarlıklarını nasıl geliştirdiklerine odaklanmamışlardır. Bu yüzden bu çalışma, Türk İngilizce öğretmenlerinin ölçme değerlendirme okuryazarlıklarını nasıl geliştirdiklerini bulmayı amaçlar. Çalışma, küçük ölçekli bir nitel çalışma olarak tasarlanmıştır. Bir Türk üniversitesinde çalışan sekiz İngilizce öğretmeni bu çalışmaya katılmıştır. Yarı-yapılandırılmış kişisel görüşmeler ve odak grup görüşmesi veri toplamak için kullanılmıştır. Çalışma; katılımcı kontrolü, üçgenleme ve yoğun anlatım kullanılarak inandırıcılığı artırılmıştır. Toplanan veri, içerik analizi kullanılarak analiz edilmiştir. Çalışmanın bulguları; önceki ölçme değerlendirme deneyiminin, ölçme değerlendirme eğitiminin ve kişisel gelişimin katılımcı İngilizce öğretmenlerinin yabancı dilde ölçme değerlendirme okuryazarlıklarının gelişimini etkilediklerini göstermiştir. Katılımcılarının; meslektaş değerlendirmeleri yaparak, teorik bilgilerini uygulamada kullanarak ve öğrencilerini ölçerek değerlendirerek kazandıkları deyimle kendilerini geliştirmelerinin, ölçme değerlendirme okuryazarlığının gelişiminde en etkili olan faktör olduğu bulunmuştur.

Anahtar Kelimeler: *Yabancı dilde ölçme değerlendirme okuryazarlığı, gelişim, İngilizce öğretmenleri*

1. Introduction

Language assessment literacy has become a popular topic studied in language assessment and evaluation research around the world and in Turkey. It is defined as language teacher's ability which he/she uses to understand, analyze and use his/her students' assessment data to improve their language learning (Inbar-Lourie, 2008a). Like Inbar-Lourie, Lam (2015) and O'Loughlin (2013) stated that LAL is the knowledge, skills, and principles which he/she understands, acquires and masters to construct a language assessment, use it, evaluate its results and understand its impact by adopting a critical attitude. Therefore, it is the knowledge, understanding, and practices of language assessment which he/she uses to create, understand, analyze and evaluate his/her language assessments in language classes (Fulcher, 2012; Malone, 2013; Pill & Harding, 2013; Scarino; 2013). Malone

(2008) defined LAL shortly as what the language teacher is supposed to know about language testing and assessment.

As LAL is what the language teacher has to know about and understand language testing and assessment, being language-assessment-literate is an important competence for the language teacher to have. It is necessary for him/her because he/she is the agent of language assessment that integrates language assessment with language teaching to improve his/her students' language learning (Malone, 2013; Rea-Dickins, 2004). Therefore, he/she is supposed to deal with every assessment-related activity in his/her classes (Rea-Dickins, 2004) including developing and preparing an exam, scoring and administering it, and interpreting its results (Alas & Liiv, 2014; Boyd, 2015; Davison & Leung, 2009; Newfields, 2006; Pill & Harding, 2013). It is also essential for him/her since he/she can meet the requirements of any change happening in language teaching methodology, educational theory and language testing and assessment (Inbar-Lourie, 2008a, 2008b). Such changes include giving a central role to his/her students, recognizing the washback effect, and knowing the social expectations from assessment and evaluation (Inbar-Lourie, 2008a). Being language-assessment-literature prepares him/her to deal with the requirements of any educational and political reform because he/she is regarded as the gate keepers of such reforms who implements reforms in his/her classes and provides visible and measurable results for accountability through assessment and evaluation (Brindley, 2008; Broadfoot, 2005; Inbar-Lourie, 2013; Malone, 2008; Rea-Dickins, 2008; Walters, 2010).

Despite LAL's significance for language teachers, different studies in Europe have shown that language teachers have a low level of LAL (Hasselgreen, Carlsen, & Helness, 2004; Vogt, Guerin, Sahinkarakas, Pavlou, Tsagari, & Afiri, 2008). Hasselgreen and her colleagues prepared the Language Testing and Assessment Questionnaire based on four different criteria: (a) classroom-focused language testing and assessment (LTA), (b) purposes of testing, and (c) content and concepts of LTA, and (d) external tests and exams. Similarly, Vogt and her colleagues used the same questionnaire by excluding the criterion (d) in the questionnaire. Both studies determined the level of LAL according to the degree of assessment training needed and received. According to the findings of both studies, the participants from different European countries including Turkey received no or little formal training about (a), (b), (c), and/or (d) (Hasselgreen et al., 2004; Vogt et al., 2008). Vogt and Tsagari (2014) found similar findings in their research in which the participants from different European countries including Turkey took part. The participant teachers in both studies had the low levels of LAL.

Similar studies were made to determine the levels of LAL of the in-service EL teachers in Turkey. In one of these studies, Büyükkarcı (2016) used Assessment

Literacy Inventory by Mertler and Campbell (2005) in which the level of LAL is determined by the number of the correct answers given by the participants. This researcher found out that the in-service EL teachers in Turkey had a low level of LAL. In addition, Hatipoğlu (2015), Öz and Atay (2017) and Şahin (2015) investigated the LAL level of the in-service EL teachers depending on the assessment training received and needed. In their studies, these researchers revealed that the LAL levels of the in-service EL teachers were low. In another study, Mede and Atay (2017) indicated that the participant in-service EL teachers in their qualitative study received incomprehensive pre-service assessment training, so the teachers were decided to have a low level of LAL.

To sum up, the literature review has indicated that the level of LAL is measured through two ways: (a) checking the knowledge of assessment and (b) finding out language assessment training received and needed. The major findings of the first group of the studies that measure LAL through checking assessment knowledge are that language teachers have assessed and evaluated their students with the low and moderate levels of LAL. None of these studies has gone beyond measuring assessment knowledge, but have mentioned that pre-service assessment training is not effective and language teachers need extra in-service assessment training to improve their LAL. The second group of the studies which focus on language assessment training to determine the level of LAL have investigated in-service and pre-service language assessment training. These studies have revealed that pre-service assessment training that language teachers have received is ineffective and language teachers needed extra assessment training to be language-assessment-literate. The studies have also mentioned that language teachers have improved their knowledge, skills and practices of assessment and evaluation on the job, but not how language teachers have contructed and developed them. Therefore, it is significant to find out how language teachers develop and construct their assessment and evaluation knowledge, skills and practices (e.g. LAL). The present study aims to reveal the development and construction of LAL by going beyond measuring the levels of LAL depending on the two ways aforementioned. In accordance with this purpose, the present research has tried to answer the research question below:

1. How do Turkish EL instructors construct and develop their knowledge, skills and practices of assessment and evalution?

2. Methodology

2.1. Research Design

A qualitative study helps to understand the underlying reasons, opinions, and motivations related to the phenomenon under investigation, so it provides insights

about the phenomenon (Creswell, 2007; Dörnyei, 2011). Therefore, this research was designed as a small scale qualitative study.

2.2. Participants

Eight Turkish EL teachers were chosen through criterion sampling because it enables to find participants who experience similar things about the phenomenon under investigation as a result of the criteria prepared to select participants (Creswell, 2007; Dörnyei, 2011). The participants in this study met the seven criteria determined by the researchers depending on the standards of assessment literacy: (a) choosing their own assessment methods, (b) developing their own assessment, (c) administering their own assessment, scoring it, and interpreting its results, (d) communicating its results, (e) making decisions about their students, curriculum, and instruction, (f) developing valid grading systems, and (g) dealing with the illegal, unethical, and unprofessional uses of their assessment (American Federation of Teachers (AFT), National Council on Measurement in Education (NCME), & National Education Association (NEA), 1990).

Table 1: Demographic Information about the Participants

Participant	Gender	Age	Experience	BA	Weekly Teaching Hours	Number of Students
1	Female	28	5	EL	21	More than 200 students
2	Female	35	10	ACL	21	More than 300 students
3	Male	30	7	ELL	27	More than 200 students
4	Male	28	7	ELT	21	More than 200
5	Male	28	5	ELT	24	More than 200 students
6	Male	35	7	ELT	15	More than 200 students
7	Male	30	9	ELT	21	More than 250
8	Male	30	9	ELL	24	More than 300 students

As Table 1 indicates, two female and six male teachers participated into the research. Their ages were between 28 and 35. They had to teach more than 20 hours every week on average and had more than 200 students in total in their classes. One of the participants graduated from the English linguistics (EL) department of a Turkish university, another participant graduated from the American culture and literature (ACL) department, and two other participants graduated from the English language and literature (ELL) department. The rest were the graduates of the English language teaching (ELT) departments.

2.3. Data Collection Tools

A focus group discussion and semi-structured individual interviews were used to collect data in this research.

> **1. Semi-structured individual interviews:** Each participant was first interviewed about how they developed themselves in language assessment and evaluation. The second interviews were used for member check, that is, each participant was asked to check the transcriptions of the first interviews and focus group discussion and to verify the findings of their content analyses.

> **2. Focus group discussion:** Seven of the participants took part in a focus group discussion to triangulate the data collected from the first semi-structured individual interviews.

2.4. Trustworthiness

Triangulation, member checks, and thick description were used in this research to make it trustworhty (Lincoln & Guba, 1985). First, the data obtained from the first interviews were triangulated with the focus group discussion. Second, each participant checked the transcriptions and analyses of the first interviews and focus group discussion in the second interviews, so they ensured that the transcriptions reflected what they wanted to explain and verified the researchers' inferences about them. Finally, the findings were described thickly by the researchers.

2.5. Data Collection Procedure

A legal permission was taken from the university before collecting data. The teachers working in the Academic English Deparament of a Turkish foundation university were informed about the research and asked whether they wanted to participate the research. Eight of them volunteered to take part in the research. The consent of each participant was taken orally and in a written form before each data collection. The participants were first interviewed individually about how they developed themselves in language assessment and evaluation. Second, seven of the participants joined a focus group discussion focusing on the same question that the first interviews concentrated on. Third, the researchers transcriped the first interviews and focus group discussions and then content-analyzed them. After the transcriptions and analyses of the collected data, the researchers made the second interviews with each participant individually to be sure that their transcriptions reflected what the participants wanted to explain and that their inferences about each participant were verified by the participants.

2.6. Data Analysis

The data collected from the focus group dicsussion and first semi-structured individual interviews were content-analyzed by following the four ways that Yıldırım and Şimşek (2013) suggested for content analysis in Figure 1.

Figure 1: Data analysis procedure
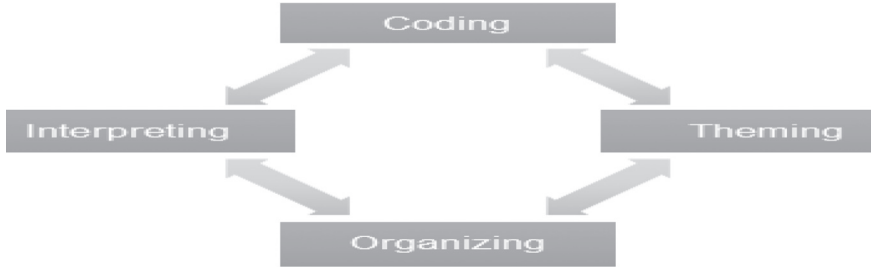
As Figure 1 shows, the researchers categorized the data into meaningful units and conceptualized what these meaningful units were by giving codes which explained the relationships in each meaningful unit. Meanwhile, they read the data many times to code and used the codes derived from the data to name them. After preparing a code list, they found the themes which covered the codes in the list by sorting out the similarities and differences among the codes, so they categorized the codes by placing the similar ones into a theme and explained the relationships among them. As a result of coding and theming, they organized and described the data with the excerpts taken from the first interviews and focus group discussion. The data collected were presented by relating them to each other without adding the researchers' comments or interpretations to the analysis. They interpreted the data without conflicting with the description of the data in the end. Then, they made explanations in order to make the data meaningful, to make logical conclusions from the findings, to reveal reason and result relationship and to show the importance of the findings.

3. Findings

The collected data were content analyzed depending on the themes and codes developd and shown in Table 2.

Table 2: The Themes and Codes Derived to Analyze the Data Collected

1. Previous assessment experience
a. Different types of assessment used in studentship
b. Previous teachers
2. Assessment training
a. Pre-service assessment training
b. Pedagogical formation training
c. In-service assessment training
3. Self-improvement
a. Peer interaction
b. Practicing theoretical information
c. Assessing and evaluating

3.1. Previous Assessment Experience

The first theme is *previous assessment experience* which every participant gained as a result of different assessment methods used to assess and evaluate them when they were students. In addition, the participants' teachers affected the way they improved their assessment and evaluation.

The use of different assessment methods in assessing and evaluating the participants when they were students affected each participant's assessment and evaluation knowledge in several ways. First of all, the methods led to the belief that if an assessment encourages a student to think, study, and produce something by using what he/she has learned, and promotes his/her self-confidence, it is the correct assessment tool for all of the participants. Besides, according to Participant 9, the assessment methods like open-ended and multiple-choice exams which his teachers had used to assess him enabled him have a job, so he thought that such assessment methods were the correct assessment methods to be used in assessing and evaluating his students. Like him, Participant 8 emphasized that the assessment methods used when he was a student enabled him to improve his language skills and become an effective teacher, so he considered that those methods like presentation, open-ended exam, and multiple-choice exams were the correct assessment tools. On the other hand, the assessment methods used when Participants 1, 2, and 3 were students affected the way they studied negatively, so they had negative attitudes toward the assessment methods like multiple-choice which caused them to lack self-confidence and self-assessment. Therefore, Participant 3 did not support using assessment as a punishment in his present classes.

In addition to different assessment methods, some of the participants were also affected by *their previous teachers* in developing their assessment and evaluation knowledge. For example, Participant 8 told, "I follow what my high school teachers did, so I do what I observed in assessing my students" and Participant 9 explained, "I thought about how the exams were prepared by our teachers in the past. I use them in my assessment as I think they are the correct assessment tools". Those participants remembered how their previous teachers prepared their questions to assess what they taught and how the participants studied for their previous teachers' exams. They also remembered that their teachers' exams helped them to find out what they learnt and did not learn.

3.2. Assessment Training

The second theme is *assessment training* which every participant received as *pre-service, pedagogical formation, and/or in-service assessment training*.

Participants 4, 5, 6, and 7 received *pre-service language assessment training* because they graduated from the ELT departments of different Turkish universi-

ties. Language assessment and evaluation course is obligatory in the ELT departments. Participants 4, 5, and 6 found their pre-service training effective. For instance, Participant 4 stated, "it [pre-service assessment training] has an effective. Its most clear effect is on what I do in terms of assessment and evaluation". However, he could not remember any theoretical information related to assessment and evaluation because he said, "a lot of years have passed since I learned those things at university". Similarly, Participant 5 considered his pre-service assessment training effective because he benefitted from what he learned in his pre-service training when he decided what and how to assess and how to prepare and organize his exams. The excerpt below clearly indicates this:

> Participant 5 (First interview): Of course, it has. For example, when I prepare an exam for students, no matter what it is, it helps me to determine the criteria like what I should pay attention to, what my goal is, and what I should assess because I learned them in my pre-service training. Besides, I learned how to prepare questions, what I should pay attention to in preparing questions, and how to format an exam paper. We learned them in our pre-service training.

Participant 6 also thought that his pre-service training was effective because his training taught him the ways to make an exam valid and reliable which he still used in assessment and evaluation like Participants 4 and 5:

> Participant 6 (First interview): ... we took an assessment and evaluation course. We prepared our tests and checked their validity and reliability. For example, I chose a topic in English. I developed a test after I taught it. It is effective. For example, can my exam which has 50 questions cover what I have taught in my classes? I check it. How are the questions going to be scored? I prepare it accordingly. It also includes preparing questions suitable to the level of students.

However, Participant 7 believed that his pre-service training was not effective because it was too theoretical and lacked practice.

On the other hand, Participants 1, 2, 3, and 8 did not take any pre-service assessment training because of the departments they graduated from. Instead, they took *an assessment course in pedagogical formation training* which they had to receive in order to be EL teachers. All of these participants did not find this course effective. Participant 1 thought that course was very theoretical for her to understand, and the course teacher's attitude toward the course was negative; therefore, she was disinterested and disengaged:

Participant 1 (First interview): What can we attribute this to? In order to take the certificate of teaching, I had to take some courses unfortunately in my pedagogical formation training. In my opinion, assessment and evaluation course was something that consisted of numerical things. I remember that I failed in this course because presentations were made and composed of theoretical knowledge and numerical values. The course teacher did not pay enough attention to our learning. As a result, I was not interested and engaged in the course. I think it was not attached enough importance.

Like Participant 1, Participant 8 also did not believe that the course in his pedagogical formation training affected his assessment and evaluation practices because he said, "the course was not taught seriously". Besides, Participant 3 thought that the course teacher did not pay enough attention to that course and give him and his friends opportunities to practice what they learned as understood from the quotation:

Participant 3 (First interview): I want to say frankly that it is a serious problem. That is, we knew that the teacher of assessment and evaluation course did not provide us with an opportunity to practice what we learned and did not pay enough attention to the course. Therefore, we should not think that we can expect a student to have the expectation that what he has learned will be useful in an environment if the teacher does not give importance to assessment and evaluation. Therefore, I had trouble in this course.

In addition to pre-service and pedagogical formation training, only Participants 1, 3, 4, and 6 received *in-service assessment training*. Participants 3 and 6 found it effective though Participant 6 could not remember its specific effect on his assessment practices. Participant 3 learnt how different types of assessment methods required students to use their different capacities as a result of his in-service training. Yet, Participants 1 and 4 did not believe that the training was effective. Participant 4 thought that he could not remember anything related to the training.

3.3. Self-improvement

The third them is *self-improvement* in assessment and evaluation as a result of *peer interaction, practicing theoretical information, and assessing and evaluating students*.

Participants 4, 5, 7, and 8 told that they improved their assessment and evaluation knowledge through *peer interaction*. Participant 4 told that he interacted and collaborated with his colleagues in assessing his students. They gave feedback to

each other and also received feedback from one another in terms of the level and quality of their exams. He believed that peer feedback and assessment improved his assessment and evaluation. To illustrate:

> Participant 4 (First interview): In fact, I did not attend any seminar related to any course and did not read any book and article about assessment and evaluation, but I have always prepared exams with my colleagues during seven years. We have made discussions and exchanged ideas about the difficulty level of our exams, the averages of our classes, and the quality of our exam questions. On the other hand, I did not do anything willingly and consciously to improve myself in assessment and evaluation.

Similarly, Participant 7 said that he shared his questions with his colleagues. They peer-assessed their exams and gave feedback to each other, which he believed helped to improve his assessment and evaluation practices. To indicate:

> Participant 7 (First interview): Now, I give my questions to my colleagues in my department. They check my questions and give feedback about them in terms of their difficulty level and relatedness to the content coverage of the exam. This also contributes to my improvement.

Similarly, Participant 5 mentioned that he worked in the Testing Office of the university. While working there, he and his colleagues checked their exams, gave and received feedback to and from each other about the exams, and made necessary changes in the exams according to peer feedback which was usually on validity and reliability, as the excerpt below clearly shows:

> Participant 5 (First interview): … but I worked in the Testing Office for almost one year. I can say that I improved my assessment and evaluation there thanks to my colleagues. For example, if I made a mistake, it was corrected by someone, and someone else's mistake was corrected by me. We were not officially trained in testing and assessment, but we learned from each other.

Participant 8 observed his colleagues when they prepared their exams. He evaluated his observations and decided to use some of the things his colleagues did:

> Participant 8 (First interview): I have completely done everything depending on my observation and experiences related to choosing assessment tools. I have im-

proved myself in an old schooled way in terms of education and training. That is, I observed my colleagues in terms of what they did. By looking at them, I developed my own way of assessment and evaluation. For example, you wonder something when your friends prepare their exam, so you observe what and how they do. You like some of the ways. They may do something different in their exams. For example, I did not use matching a lot in my exams, but when I saw that my friends used matching, I started to use it.

In addition to peer interaction, Participants 2, 4, 5, and 6 developed their assessment and evaluation knowledge by *practicing theoretical knowledge* in which they integrated their theoretical knowledge with their assessment and evaluation practices. For example, Participant 5 stated, "In fact, I had a chance to use the theoretical information that I learned at school for almost one year", and Participant 2 explained, "I attended a CELTA program abroad. I improved myself by using what I learned there; practicing my learning, and cooperating with my colleagues."

Moreover, all of the participants told that *assessing and evaluating their students* is the most important and effective way of developing their assessment and evaluation knowledge. That is, they learned how to assess and evaluate by assessing and evaluating their students. For example, Participant 6 told that he improved himself on the job by gaining experience as the excerpt below illustrates:

Participant 6 (First interview): In fact, a teacher learns how to teach and how to assess while he is working like others who learn their jobs when they work. In fact, a teacher develops his own way depending on his teaching conditions by adapting what was told in his university courses and in-service workshops. Namely, he forms his own way of teaching by keeping what he has learned in mind and combining them with his teaching conditions.

Similarly, Participant 3 improved his assessment and evaluation practices by gaining experience through assessing and evaluating his students:

Participant 3 (First interview): I improved myself especially by gaining experience in assessment and evaluation. As you know, the most troublesome part at some universities in Turkey is the theoretical part. That is, they teach theories to students, but do not give them enough opportunities to practice. The most problematic thing for us is that we know the theory, but cannot put it into practice. Therefore, if you ask a student in any discipline, he says "I did not learn anything at the university. I learned some things in my professional life". Similarly, I learned thanks to my experiences in my working life.

All of the participants gained experience by making assessment and evaluation on their own. They told that, through experience, they developed their assessment knowledge and improved their awareness in practice by self-assessing their experiences and also transferring and adapting what they learned to their new teaching contexts.

4. Discussion

According to the the findings of the study, the use and preparation of different assessment methods to assess the participants when they were students constituted their previous assessment experience. This old experience has been found to cause the participants to form different beliefs about different assessment methods depending on the effects of assessment methods, which is a part of teachers' assessment and evaluation knowledge. This result adds a new perspective to the finding in the literature that language teachers develop their assessment and evaluation on the job (Hasselgreen et al., 2004; Vogt et al., 2008; Vogt & Tsagari, 2014). It is because assessment beliefs can help to understand the assessment and evaluation knowledge related to why EL teachers consider some assessment methods as the correct assessment tools, how they determine the efficiency of such methods, and why they avoid using other assessment methods.

The findings of the study have shown that there are three sources of language assessment training for the participant teachers: pre-service, pedagogical formation and in-service assessment training. As the four participants graduated from the ELT departments, they received pre-service assessment training, which most participants in this group found effective in their assessment and evaluation practices. This finding conflicts with much research in the literature which claimed that the participant language teachers in these studies had the low levels of LAL because they received no or little pre-service assessment training (Hasselgreen et al., 2004; Hatipoğlu, 2015; Mede & Atay, 2017; Öz & Atay, 2017; Şahin, 2015; Vogt et al., 2008; Vogt & Tsagari, 2014). Therefore, pre-service language assessment training seems to contribute to the development of LAL. The second source is pedagogical formation assessment training. Four participants received such training because the departments (e.g. ACL, ELL and EL) they graduated from did not include any language assessment course in their curricula. These participants thought that it was very ineffective and incomprehensive for them. Pedagogical formation assessment training can be considered as a type of pre-service training since students of ACL, ELL and EL generally receive education courses including assessment course before they start to work as EL teachers in Turkey. Considering this assumption and the participants' thought about pedagogical formation, this finding of the present research proves the findings of much research in the literature which revealed that

pre-service assessment training has no or little effect on the development and construction of LAL (Hasselgreen et al., 2004; Hatipoğlu, 2015; Hatipoğlu & Erçetin, 2016; Mede & Atay, 2017; Öz & Atay, 2017; Şahin, 2015; Vogt et al., 2008; Vogt & Tsagari, 2014). Therefore, pedagogical formation assessment training can be thought not to affect the development and construction of LAL. The last source of language assessment training is in-service training which four participants received and the rest did not receive. This result has revealed that language teachers lacked in-service language assessment training. In addition, two participants did not believe that such training affected them in assessment and evaluation. These two results corroborate the findings of different researchers (Hasselgreen et al., 2004; Hatipoğlu, 2015; Öz & Atay, 2017; Şahin, 2015; Vogt et al., 2008; Vogt & Tsagari, 2014) that revealed that langauge teachers received no or little assessment training (e.g., in-service training and pre-service).

As the findings in the present study have pointed out, pre-service assessment training can help EL teachers to (a) understand how to prepare exams by focusing on what to pay attention in preparing their exams, about the purpose of assessment and how to prepare the layout of the exams, (b) adjust the difficulty levels of their questions according to the levels of their students, and (c) make exams valid and reliable in terms of the contribution of pre-service assessment training to the development and construction of LAL. On the other hand, the reasons of inefficiency of pre-service training including pedagogical formation are listed as it may be too statistical and lack practice, and because the teachers of the courses may not give enough importance to such assessment courses, which is consistent with the findings of several researchers (e.g., Lomax, 1996; Rogier, 2014; Stiggins, 1991).

The third and most important finding of the study is that the participants developed and constructed LAL on the job through self-assessment, which proves the findings of much research in the literature (Hasselgreen et al., 2004; Hatipoğlu, 2015; Mede & Atay, 2017; Öz & Atay, 2017; Şahin, 2015; Vogt et al., 2008; Vogt & Tsagari, 2014). Self-improvement benefits from different sources. The first source is peer interaction. It enables EL teachers to learn from each other because they give and receive feedback about their exams to and from each other, so they have a change to see how their colleagues prepare their questions, what kind of problems their and their colleagues exams have, and how such problems can be overcome, which is in line with what Munoz, Palacio and Escobar (2012), Scarino (2013) and Tahmasbi (2014) found out in their studies. In addition, working as a teacher gives EL teachers opportunities to practice what they learn theoretically because EL teachers can put theory in practice, understand how it works and recognize its strengths and weaknesses (Izci & Siegel, 2014; Stiggins, 1995). The third way is assessing and evaluating students. It seems to be the most effective

source in developing and constructing LAL through self-assessment because it helps EL teachers to adopt a more critical attitude in assessment and evaluation a lot. They learn how to deal with different assessment situations when they assess and evaluate the students. They also learn how to use what they have learnt in in-service and pre-service assessment training as well as from peer interaction and their previous assessment experiences. Being critical in assessment and evaluation enables them to self-assess their assessments so that they can improve the assessments by finding out and dealing with the weaknesses in the exams (AFT et al., 1990; Tulgar, 2017). To conclude, self-improvement fed with different resources can help EL teachers to develop and construct their LAL particularly by giving EL teachers opportunities to adopt a critical attitude, which is a very significant part of LAL (Inbar-Lourie, 2008a; Lam, 2015; O'Loughlin; 2013). This finding has also contributed to the findings of several studies in the literature (Hasselgreen et al., 2004; Hatipoğlu, 2015; Mede & Atay, 2017; Öz & Atay, 2017; Şahin, 2015; Vogt et al., 2008; Vogt & Tsagari, 2014) by explaining how language teachers develop and construct LAL (e.g., assessment knowledge, skills and practices.

5. Conclusion

Different studies in the literature point out that the in-service Turkish EL teachers have low levels of LAL (Büyükkarcı, 2016; Hasselgreen et al., 2004; Hatipoğlu, 2015; Mede & Atay, 2017; Öz & Atay, 2017; Şahin, 2015; Vogt et al., 2008; Vogt & Tsagari, 2014), which means they may not have sufficient assessment and evaluation knowledge when they assess and evaluate students in the class. This also means that they should develop themselves in assessment and evaluation, but none of the studies mentioned in this study has indicated how this development occurred.

This study has helped to understand how LAL is developed and constructed by EL teachers by explaining the finding in the literature that language teachers have developed themselves on the job (Hasselgreen et al., 2004; Hatipoğlu, 2015; Öz & Atay, 2017; Şahin, 2015; Vogt et al., 2008; Vogt & Tsagari, 2014). The study has revelaed that there are three sources of the development and construction of LAL: (a) previous assessment experience, (b) assessment training and (c) self-improvement. According to the study, previous assessment experience can result in forming different beliefs about assessment methods, which is an important part of EL teachers' assessment knowledge. The study has also revealed that the effects of assessment training on EL teachers can change depending on the departments EL teachers graduate from. Pre-service assessment training can be effective in developing the assessment knowledge of the EL teachers graduating from ELT departments. However, pedagogical formation assessment training cannot be as

effective as pre-service assessment training because it may be too theoretical, lack practice, and may not be given enough importance by the course teachers. Besides, the study has pointed out that EL teachers can improve themselves in assessment and evaluation through peer interaction, practicing theoretical information and gaining experience by assessing and evaluating students, all of which are the sources of self-improvement. Self-improvement can teach them how to deal with different assessment situations and how to use what they have learned from different sources by adopting a critical attitude.

Further studies can be made on the content of assessment training in pedagogical formation training. The findings of this study can be used to develop an in-service assessment training which may give opportunities to practice theoretical information by concentrating on peer interaction, so further studies can be conducted about its effects on EL teachers' assessment knowledge and practices. Besides, how self-improvement in assessment and evaluation influences EL teachers' assessment knowledge, beliefs, and practices can be investigated in future studies.

REFERENCES

ALAS, E. & LIIV, S. (2014). Assessment literacy of national examination interviewers and rater-experience with the CEFR. *EESTI RAKENDUSLINGVISTIKA ÜHINGU AASTARAAMAT, 10*, 7-22.

AMERICAN FEDERATION OF TEACHERS, NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION, & NATIONAL EDUCATION ASSOCIATION. (1990). *Standards for teacher competence in educational assessment of students*. Retrieved from http://buros.org/standards-teacher-competence-educational-assessment-students

BRINDLEY, G. (2008). Educational reform and language testing. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed.) (pp. 365-378). Springer Science+Business Media LLC.

BROADFOOT, P. M. (2005). Dark alleys and blind bends: Testing the language of learning. *Language Testing, 22*(2), 123-141. doi: 10.1191/0265532205lt302oa

BÜYÜKKARCI, K. (2016). Identifying the areas for English language teacher development: A study of assessment literacy. *Pegem Eğitim ve Öğretim Dergisi, 6*(3), 333-346.

CRESWELL, J. W. (2007). *Qualitative inquiry & research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, California, the United States of America: Sage Publications.

DAVISON, C. & LEUNG, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly, 43*(3), 393-415.

DÖRNYEI, Z. (2011). *Research methods in applied linguistics*. Oxford, UK: Oxford University Press.

FULCHER, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly, 9*(2), 113-132. doi: 10.1080/15434303.2011.642041

HASSELGREEN, A., CARLSEN, C., & HELNESS, H. (2004). *European survey of language testing and assessment needs. Report: Part one – general findings*. Retrieved fromhttp://www.ealta.eu.org/documents/resources/survey-report-pt1.pdf

HATIPOĞLU, Ç. (2015, May). *Diversity in language testing and assessment literacy of language teachers in Turkey*. Paper presented at the 3rd ULEAD Congress, International Congress on Applied Linguistics: Current Issues in Applied Linguistics, Çanakkale, Turkey.

HATIPOĞLU, Ç. & ERÇETIN, G. (2016). Türkiye'de yabancı dilde ölçme ve değerlendirme eğitiminin geçmişi ve bugünü. In S. Akcan & Y. Bayyurt (Eds), *3. ulusal yabancı dil eğitimi kurultayı: Türkiye'deki yabancı dil eğitimi üzerine görüş ve düşünceler 23-24 Ekim 2014, konferanstan seçkiler* (pp. 72-89). İstanbul, Turkey: Boğaziçi University.

INBAR-LOURIE, O. (2008a). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing, 25*(3), 385-402. doi: 10.1177/0265532208090158

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. (2013). Language assessment literacy. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1-8). Blackwell Publishing Ltd.

IZCI, K. & SIEGEL, M. (2014, April). *Investigating high school chemistry teachers' assessment literacy in theory and practice*. Paper presented at the annual meeting of the American Educational Research Association, Philadelphia, PA.

LAM, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing, 32*(2), 169-197. doi: 10.1177/0265532214554321

LINCOLN, Y. S. & GUBA, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.

LOMAX, R. G. (1996). On becoming assessment literate: An initial look at preservice teachers' beliefs and practices. *The Teacher Educator, 31*(4), 292-303. doi: 10.1080/08878739609555122

MALONE, M. E. (2008). Training in language assessment. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed.) (pp. 225-239). Springer Science+Business Media LLC.

MALONE, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing, 30*(3), 329-344. doi: 10.1177/0265532213480129

MEDE, E. & ATAY, D. (2017). English language teachers' assessment literacy: The Turkish context. *Ankara Üniversitesi TÖMER Dil Dergisi, 168*(1), 43-60.

MERTLER, C. A. & CAMPBELL, C. (2005, April). *Measuring teachers' knowledge & application of classroom assessment concepts: Development of the Assessment Literacy Inventory*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

MUNOZ, A. P., PALACIO, M., & ESCOBAR, L. (2012). Teachers' beliefs about assessment in an EFL context in Colombia. *PROFILE, 14*(1), 143-158.

NEWFIELDS, T. (2006). *Teacher development and assessment literacy*. Paper presented at

the Proceedings of the 5th Annual JALT Pan-SIG Conference, Shizuoka, Japan (pp. 48-73). Tokai University College of Marine Science.

O'LOUGHLIN, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing, 30*(3), 363-380. doi: 10.1177/0265532213480336

ÖZ, S. & ATAY, D. (2017). Turkish EFL instructors'in-class language assessment literacy: Perceptions and practices. *ELT Research Journal, 6*(1), 25-44.

PILL, J. & HARDING, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing, 30*(3), 381-402. doi: 10.1177/0265532213480337

REA-DICKINS, P. (2004). Understanding teachers as agents of assessment. *Language Testing, 21*(3), 249-258.

REA-DICKINS, P. (2008). Classroom-based language assessment. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed.) (pp. 257-271). Springer Science+Business Media LLC.

ROGIER, D. (2014). Assessment literacy: Building a base for better teaching and learning. *English Teaching Forum, 3*, 2-13.

ŞAHIN, S. (2015, May). *Language testing and assessment (LTA) literacy of high school English language teachers in Turkey*. Paper presented at the 3rd ULEAD Congress, International Congress on Applied Linguistics: Current Issues in Applied Linguistics, Çanakkale, Turkey.

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. doi: 10.1177/0265532213480128

STIGGINS, R. J. (1991). Assessment literacy. *The Phi Delta Kappan, 72*(7), 534-539.

STIGGINS, R. J. (1995). Assessment literacy for the 21st century. *The Phi Delta Kappan, 77*(3), 238-245.

TAHMASBI, S. (2014). Scaffolding and artifacts. *Middle-East Journal of Scientific Research, 19*(10), 1378-1387.

TULGAR, A. T. (2017). Selfie@ssessment as an alternative form of self-assessment at undergraduate level in higher education. *Journal of Language and Linguistics Studies, 13*(1), 321-335.

VOGT, K., GUERIN, E., SAHINKARAKAS, S., PAVLOU, P., TSAGARI, D., & AFIRI, Q. (2008). *Assessment literacy of foreign language teachers in Europe – current trends and future perspectives*. Paper presented at the 5th EALTA Conference, Athens, Greece.

VOGT, K. & TSAGARI, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly, 11*, 374-402. doi: 10.1080/15434303.2014.960046

WALTERS, F. S. (2010). Cultivating assessment literacy: Standards evaluation through language-test specification reverse engineering. *Language Assessment Quarterly, 7*(4), 317-342. doi: 10.1080/15434303.2010.516042

YILDIRIM, A. & ŞİMŞEK, H. (2013). *Sosyal bilimlerde nitel araştırma yöntemleri* (9th ed.). Ankara: Seçkin Yayıncılık.