EXPLORING FOREIGN LANGUAGE ASSESSMENT LITERACY TRAINING NEEDS OF PRE-SERVICE ENGLISH LANGUAGE TEACHERS[1]

(Research article)

(corresponding author)
Aylin Sevimel-Sahin https://orcid.org/0000-0003-2279-510X
ELT Department, Anadolu University, Eskisehir, Turkey
aylinsevimel@anadolu.edu.tr

Gonca Subasi https://orcid.org/0000-0001-7049-5940
ELT Department, Anadolu University, Eskisehir, Turkey
goncas@anadolu.edu.tr

Biodata:

Dr. Aylin Sevimel-Sahin is currently working in ELT Department, Faculty of Education, Anadolu University, Eskisehir, Turkey. Her research interests are ELT teacher education, language testing and assessment, teaching practice, affective domain, use of digital tools in language education, and research methodology.

Gonca Subaşı is an assistant professor doctor in the ELT Department at Anadolu University, Turkey. Her research interests include teaching writing skills, vocabulary acquisition, affective factors in language teaching, language testing and evaluation, and language teacher education.

---

[1] This paper is based on the doctoral dissertation titled 'Exploring foreign language assessment literacy of pre-service English language teachers' of the first author.

EXPLORING FOREIGN LANGUAGE ASSESSMENT LITERACY TRAINING NEEDS OF PRE-SERVICE ENGLISH LANGUAGE TEACHERS

Aylin Sevimel-Sahin
aylinsevimel@anadolu.edu.tr

Gonca Subasi
goncas@anadolu.edu.tr

Abstract

Language teachers are expected to be language assessment literate (LAL) - the ability to use assessment knowledge effectively in own contexts - to guide instruction and improve learning. However, research demonstrated that most English teachers had challenges in assessment practices and needed further training. Then, this study aims to explore LAL training needs of pre-service English language teachers through a course-based training in teacher education. The participants as pre-service English language teachers responded to open-ended questions two times: before the course (N=164) and after the course (N=146). The content analysis of their responses yielded two major themes: conceptual dimension and course-related issues. Although most found the course beneficial and sufficient for LAL training, certain training needs were revealed. For example, the course was too theoretical and lacked practice. The topics were testing-oriented and other topics such as formative assessment, evaluation skills, integrated testing, ethical issues, and contextual factors were not included. Therefore, some implications were discussed to revise the course for a better LAL training at the teacher education phase.

Keywords: English language teaching, language assessment literacy, language testing, pre-service teacher, teacher education

1. Introduction

Language assessment is an integral part of language teaching. Without effective assessment practices, teaching would be deficient in responding to the needs of its educational context. It is because language assessment is helpful to find out what and how much is learned, what kind of learning difficulties there are, whether the goals are accomplished, and whether the style or method of teaching as well as the materials used are effective (Herrera & Macias, 2015; Hidri, 2018; Rogier, 2014). That is, language assessment is useful to get feedback about the quality of learning and instruction, and it can be used with different purposes or functions to make progress in language learning. According to the outcomes of assessment, language teachers can organize their teaching better and make necessary decisions to promote and reinforce learning while language learners can detect their strengths and weaknesses to regulate their own performance (Bachman, 2005; Rea-Dickins, 2004). Therefore, language assessment becomes a motivating focal point both for teachers and learners in education.

Language teachers are acknowledged as the main stakeholders in language learning and instruction (Giraldo, 2018) because they have the primary responsibility for planning and monitoring teaching/learning processes. Hence, they need to be qualified in certain competencies: language proficiency, language teaching and language assessment. Among these competencies, language assessment is crucial to guide the educational decisions of

teachers since such decisions can have both a positive and negative impact on teaching/learning itself (Inbar-Lourie, 2013). For this reason, teachers need to be competent evaluators to be able to fulfill the needs of their own teaching contexts. However, to perform effective assessment procedures, they need to possess a certain competence, namely, *language assessment literacy (LAL)*, as their professional competence.

LAL basically refers to the familiarity with and the ability to carry out all language assessment related issues. In other words, teachers need to know what testing, assessment and evaluation are, what kind of techniques can be used effectively for which reason or purpose at which time according to own contexts, how they can be designed and administered efficiently, and how they can be interpreted to support and monitor language development (Inbar-Lourie, 2017; Jeong, 2013; Pill & Harding, 2013). In addition, not only the knowledge of and familiarity with the assessment concepts and procedures but also the ability and skills are covered in the definition of LAL since teachers practice assessment in their classrooms by developing, administering, using, interpreting, analyzing, and evaluating language tests and assessment tools in order to monitor the progress of language learners and give constructive feedback in line with the results (Lam, 2014; Vogt & Tsagari, 2014). More extensively, Fulcher (2012) put forth the definition of LAL as:

> The knowledge, skills and abilities required to design, develop, maintain or evaluate, large-scale standardized and/or classroom-based tests, familiarity with test processes, and awareness of principles and concepts that guide and underpin practice, including ethics and codes of practice. The ability to place knowledge, skills, processes, principles and concepts within wider historical, social, political and philosophical frameworks in order to understand why practices have arisen as they have, and to evaluate the role and impact of testing on society, institutions, and individuals (p. 125).

Then, this definition is assumed to be too expanded in terms of its coverage by reflecting the understanding of LAL in all aspects. However, in this study, foreign language assessment literacy (FLAL) term has been preferred to directly represent English as a foreign language (EFL) context.

Depending on such definitions, some researchers discussed the possible dimensions to present a conceptual framework. For example, Davies (2008) put forward in relation to language testing that LAL has three components as *knowledge* (the understanding of language description, setting and measurement), *skills* (the training of developing tests), and *principles* (the use of tests appropriately and ethically). Similarly, Inbar-Lourie (2008) asserted that LAL is made up of *what* (the language trait to be measured), *how* (the execution of assessment), and *why* (the reasons behind assessment decisions and choices). Furthermore, Fulcher (2012) proposed a three-dimensional model of LAL: *practices* (knowledge, skills, and abilities), *principles* (processes, principles and concepts), and *contexts* (historical, social, political, and philosophical frameworks: origins, reasons and impacts). Regarding contexts, Hill (2017), Inbar-Lourie (2013), Stabler-Havener (2018), and Yan, Zhang, and Fan (2018) also underlined the significance of instructional contexts in that the features of local teaching contexts such as the needs and expectations of own environments, sociocultural values, educational and institutional policies are influential in undertaking suitable assessment practices. Considering such arguments, language assessment literate teachers have already been found to have common characteristics such as the knowledge and skills of language assessment theories and practices in line with the instruction, the skill of designing, administering and evaluating assessment tools, the ability to carry out assessment procedures appropriately and ethically related to the contexts, and the capability of interpreting

assessment results to sustain better teaching/learning (Berry & O'Sullivan, 2016; Bohn & Tsagari, 2021; Huang & He, 2016; Taylor, 2009). Consequently, foreign language teachers need to have such characteristics to be assessment literate for effective teaching practices as well as for cultivating language learning more.

To be language assessment literate is not easy as it seems. It makes education and training a requisite because 'teachers are not born testers' (Jin, 2010, p. 556). Lately, there is much more emphasis on LAL training due to the changing concepts and roles in foreign language education. Particularly, there has been a shift from traditional testing notion to measure learning as the product (summative) to the present conception of assessment to monitor and improve learning during the process (formative) (Csépes, 2014). This results from the effect of the changes in language teaching/learning methodologies towards a more learner-centered approach (Tsagari & Vogt, 2017). That is, teaching and assessment were considered separate earlier (Viengsang, 2016); hence, language assessment training was ignored and overshadowed by methodology education in the past. But recently, there is a growing emphasis on classroom assessment and the role of teachers as assessors because of the interplay between teaching and assessment (Scarino, 2013; Wach, 2012), which has put more responsibility on teacher education to equip future teachers with LAL competencies. Then, regarding all the LAL discussions in the literature, LAL training can encompass:

> an appropriate balance of technical know-how, practical skills, theoretical knowledge, and understanding of principles, but all firmly contextualized within a sound understanding of the role and function of assessment within education and society (Taylor, 2009, p. 27).

Nevertheless, the research into the concept of LAL is said to be new and needs to be developed (Fulcher, 2012; Inbar-Lourie, 2013) because there are few studies with pre-service teachers in English Language Teaching (ELT) education but several studies with in-service teachers. Most of the studies with teacher candidates focused on their experiences and beliefs about LAL and assessment practices in practicum. To illustrate, many teacher candidates of ELT were found to have lower levels of LAL because they received a basic level of training; thus, they could not perform any testing or assessment in teaching practice properly, especially the formative one (Howerton, 2016; Karaman & Sahin, 2014; Viengsang, 2016). Besides, they developed some misunderstandings about the assessment concepts and choices, i.e., testing was seen equal to assessment, and they learned about only summative assessment (Viengsang, 2016). Though they seemed to be ready for classroom assessment, they did not practice assessment in pre-service teacher training phase much (Komur, 2018).

Concerning LAL training content in teacher education, in spite of very few studies, the researchers investigated mostly its final effect through the given specific courses related to language assessment. For instance, some studies concluded pre-service ELT teachers were contented with their LAL training courses at their teacher education programs and found them beneficial because their views and attitudes changed positively towards assessment, and gained certain abilities related to language assessment (Hilden & Frojdendahl, 2018; Watanabe, 2011). However, most studies revealed lacking points of LAL training in teacher education. For example, pre-service ELT teachers did not have any chance to practice what they learned related to language assessment since course instructors dealt with theoretical issues more (Hatipoglu, 2010; Jin, 2010; Lam, 2014; Sariyildiz, 2018). Moreover, most syllabi of the LAL courses included only testing issues (Hatipoglu, 2015; Jin, 2010). Besides, teacher candidates had challenges in designing and administering language tests, especially according to language level (Ukrayinska, 2018) as well as in scoring, grading, evaluating, and interpreting them (Watanabe, 2011). In addition, a number of researchers revealed the

question of how well teacher education programs trained their undergraduates as future teachers in language assessment despite LAL training because many English teachers graduated with a poor ability to observe and track changes in language development through assessment and to prepare and practice assessment tools (e.g., Berger, 2012; Chen, 2005; Gan & Lam, 2020; Malone, 2008; Vogt & Tsagari, 2014). This issue is not so different in the Turkish context. Despite very few studies in language assessment, especially, about its training side, they concluded teacher training programs remain inadequate to provide the expected training in language assessment (e.g. Hatipoglu, 2010, 2015; Mede & Atay, 2017; Turk, 2018). Therefore, all those studies demonstrated that most language assessment courses given at universities were not effective for developing LAL (e.g., Giraldo & Murcia, 2018; Lam, 2014), and teacher candidates needed more training in different aspects of language assessment. Though pre-service teacher education is of utmost important to support LAL training early in learning and LAL is highly contextual rather than universal (Bohn & Tsagari, 2021; Xu & Brown, 2016), those studies only reported the general findings after the course but did not investigate what is required to improve LAL training by taking the needs of pre-service teachers into account during the training phase within their own educational context.

Considering all the arguments and findings in the literature, there is a demand to understand what teacher candidates need about LAL in relation to the course content during their training phase in their own contexts and how the course-based LAL training meet their needs throughout this training process in order to help them to be capable and competent teachers for their careers as well as to develop LAL training at the teacher education phase accordingly. Therefore, the present study aims to explore LAL training needs of pre-service ELT teachers, and also to investigate whether a course-based LAL training in teacher education is useful to develop LAL of teacher candidates. Accordingly, two research questions were addressed:

1. What are the assessment training needs of pre-service English language teachers?

2. How does 'English language testing and evaluation' course affect language assessment literacy of pre-service English language teachers?

All in all, the current study is hoped to contribute to the field of language assessment research considerably. First, it is believed to provide further insights into LAL training to meet the changing needs and expectations of pre-service ELT teachers. Second, it may reveal the ways to support the initial training on LAL by describing the present status of the course and uncovering the potential assessment needs for future guidelines to develop coursework. It is because both individual and contextual factors can lead to differences in training needs, and thereby the training content. Last, since there has been a limited number of works within this research scope, to study on LAL training may contribute to the field by revealing possible new perspectives for further implications.

2. Methodology

The present research is a qualitative study incorporating needs assessment in order to explore LAL training needs. As the name suggests, needs assessment aims at gathering information about the deficiencies or gaps about the target situation to discover the needs as well as to evaluate the status quo of any existing course, program, and the like for planning and improving the content to fulfill the needs (Royse et al., 2009). Therefore, considering the purposes of this study, needs assessment was preferred to describe, analyze, and evaluate the course-based LAL training through collecting qualitative data about the opinions and

experiences of the course-takers to reveal their perceived assessment training needs within the context of ELT teacher education.

2.1. Research Context and Participants

In Turkey, the universities with Faculty of Education offer a four-year ELT teacher education program at the undergraduate level through English instruction. There are compulsory and elective courses related to the field of ELT as well as general educational and pedagogical courses. In the last year, pre-service teachers need to complete their two-semester teaching practice before graduation.

For LAL training, teacher candidates take compulsory 'English Language Testing and Evaluation (ELTE)' course in the last term (the eight semester), and is delivered three hours a week. In the current study, the participants were the ELTE course-takers as their LAL training at a Turkish state university. The course content consists of basic terms and concepts about language assessment, the design and evaluation of different language skills testing tools, mostly in the form of classroom-based tests that are appropriate for various types of schools, language levels and age groups. When the data were collected, there were eight classes of the fourth-year ELT undergraduate students enrolled in ELTE course, and those classes were instructed by four lecturers. Since the opinions and experiences of the participants were asked two times for the purposes of the study, there were 164 pre-service ELT teachers before they began the ELTE course, and 146 course-takers after they finished the course (see Table 1). All of the participants took the course for the first time and participated in this study voluntarily through purposeful sampling.

Table 1. *The distribution of the participants (numbers)*

	Responded	Excluded	Rest (main one)
Pre-service ELT teachers (BEFORE)	168	4	164
Pre-service ELT teachers (AFTER)	147	1	146

2.2. Data Collection and Analysis

In the present study, open-ended questions were used to collect data. It is because they are helpful to get a deeper understanding of the topic by analyzing the reasons, causes, experiences, and the like compared to close-ended questions (Creswell, 2012), and also more useful to reach a larger sample compared to interviews (Gay et al., 2012). The open-ended questions were formed by the researcher via thorough analysis of the related literature, especially considering the conceptualizations suggested by Davies (2008), Fulcher (2012), Hill (2017), and Inbar-Lourie (2008). Besides, these questions were validated by three experts in the language assessment field as well as the qualitative research by checking the concept suitability, the relation with the scope of this study, and the clarity and appropriateness of the language. There were six questions asked at the beginning of the course semester and other six questions at the end of the semester. Both of the questions were different from each other: The first-round questions (before the course) were about the opinions, needs and expectations of LAL training, and the second-round questions (after the

course) were related to their opinions and experiences of the course training. The consent forms were also given for the ethical concerns.

This study was conducted in the Spring Semester of 2017-2018 Academic Year when the participants were taking the ELTE course for the first time. Before data gathering, the research ethics approval from the institutional review board was granted. Then, the open-ended questions accompanied by the consent forms were distributed to pre-service ELT teachers who were taking ELTE course at their regular class times. The first application was carried out at the beginning of the semester, and 168 pre-service teachers responded to the questions but four of them were eliminated owing to incomplete papers or repeating the course. As a result, 164 participants answered the questions for the first application. Afterwards, the second application was conducted at the end of the course semester, and 147 course-takers responded to them. However, one of them was excluded due to incomplete responses; thus, there were 146 participants for the second application (see Table 1).

To analyze the responses, content analysis was performed inductively to the collected qualitative data. For the analysis, the suggestions by Creswell (2012) and Miles, Huberman, and Saldana (2014) were taken into account. First, to get the general sense of the data, the researcher examined each response by hand for the preliminary exploratory analysis. As a result, a few main codes and themes were revealed. That first analysis was shown to a field expert to get feedback. Then, all the data were embedded into NVivo 11 Pro Program for a detailed content analysis. After coding the data chunks and labeling themes, the list of the findings was made for expert checking. Two ELTE course lecturers evaluated the findings, and some misunderstandings were clarified in the organization and the codes/themes were refined for more meaningful wording. To establish the interrater-reliability, Miles, and Huberman's (1994) intercoder-agreement was used and a high consistency between the raters was achieved (97% agreement). Finally, all the findings were compared and evaluated together to get a whole understanding of the study focus.

3. Findings

The analysis of the responses of the sample revealed two major themes and certain minor themes regarding FLAL. Since there were different questions asked for two applications (before and after), the data for each application was done separately and then, discussed together through comparison.

3.1. Before the Course

At the beginning of the semester, 164 course-takers responded to the open-ended questions. The analysis of their responses revealed two themes: *FLAL-ELTE Concept*, and *ELTE Course-related Issues* (see Figure 1). The participants commented more on ELTE course-related issues (f=1046) than FLAL-ELTE concept (f=253).
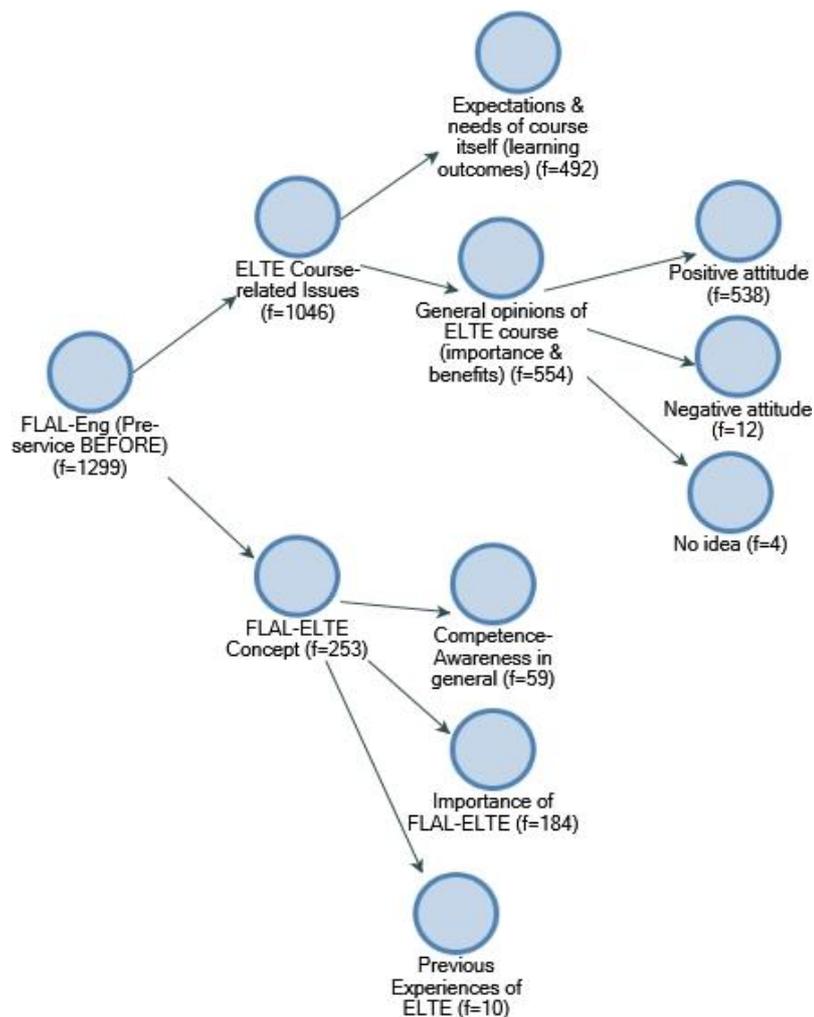
Figure 1. The thematic display of the findings from pre-service ELT teachers (before the course)

For *'FLAL-ELTE Concept',* they underlined what kind of competencies are essential for English language assessment (*f*=59), mentioned the importance of FLAL-ELTE itself (*f*=184), and shared their opinions about the previous experiences of own assessment related issues (*f*=10) (see Figure 1). Firstly, they pointed out that language teachers must know not only how to teach but also how to test since teaching and assessment are embedded in each other. Therefore, they believed teacher candidates should gain necessary competencies and skills about language assessment during their teacher *education ('Competence-Awareness in general').* For example, one of them highlighted that:

'Assessment in ELT is necessary. Being a teacher is not just about teaching. One of the most important points that every teacher has to pay attention is to prepare language exams suitable for their students. Therefore, every teacher candidate needs to be competent in language assessment.' (Before.Pre.23)

Secondly, all the participants agreed that assessment is as important as teaching itself because it is helpful to find out whether students learn and get feedback about the effectiveness of teaching *('Importance of FLAL-ELTE').* Lastly, some course-takers mentioned their previous experiences about testing as test-takers. Mostly, they complained about taking poorly prepared and wrong exams throughout their school lives, and also

criticized their previous English language teachers not knowing how to test properly (*'Previous experiences of ELTE'*).

As for '*ELTE Course-related Issues'*, the participants focused on their expectations and needs about the course itself (*f*=492) and expressed their general opinions about the course (*f*=554) (see Figure 1).

Regarding *'Expectations and needs of course'*, most participants hoped to gain more awareness about language assessment, acquire a critical eye towards language testing qualities, learn how to assess language skills and be able to effectively test and evaluate during the course period. Specifically, they expected to gain knowledge and ability of testing such as preparation, designing, and evaluating language tests according to certain criteria, all of which would be helpful for their profession. For instance, one of them stated:

'In this course, I think we will learn how to assess language learners. I expect the course will be relatively permanent, and I want to gain the knowledge that I will use in my professional teaching life.' (Before.Pre.28)

In addition, some participants indicated the ELTE course would be a revision of English teaching-methods-courses they took for three years, and they hoped to relate their methodological knowledge with assessment knowledge. Furthermore, only a few course-takers explained the backwash effect of testing and discussed negative effects of tests on learners. Unlike those statements, just a few participants only expected to enjoy and pass the course.

Considering *'General opinions of ELTE course'*, nearly all the participants had positive attitudes towards the course. They emphasized the importance of the course for their future career: it would be complementary to their teaching competence and be beneficial to change their perspectives and attitudes positively towards testing. Therefore, one of them reported that:

'I think this course should be given. It is necessary for pre-service teachers studying at faculty of education in order to learn and know how to assess language skills and the evaluation criteria for a successful and qualified professional life.' (Before.Pre.7)

Besides, some participants focused on the Turkish examination system and highlighted that exams are part of their lives. They also criticized exams are very problematic and have a great impact on test-takers. Therefore, this course was hoped to be helpful to increase their knowledge and skills of assessing language proficiency as well as to gain awareness about testing. One of them underlined that:

'In my opinion, testing and evaluation are one of the most important stages of educational process. Assessment is feedback both for teachers and students. While students have the chance to see their mistakes and learn the correct ones, teachers evaluate their teaching techniques [as a result of assessment]. Furthermore, since every stage of educational system in Turkey is dependent on exams, every teacher should know how to prepare tests. Therefore, this course is important and necessary.' (Before.Pre.58)

To sum up, the participant course-takers expected to learn the required testing knowledge and skills for their profession; hence, they believed the course would be very helpful to make them ready to assess and deal with testing issues effectively.

3.2. After the Course

At the end of the semester, 146 course-takers responded to the open-ended questions. The analysis of their responses revealed two themes: *FLAL-ELTE Concept*, and *ELTE Course-related Issues* (see Figure 2). The participants reflected on ELTE course-related issues (f=1036) more than FLAL-ELTE concept (f=617).
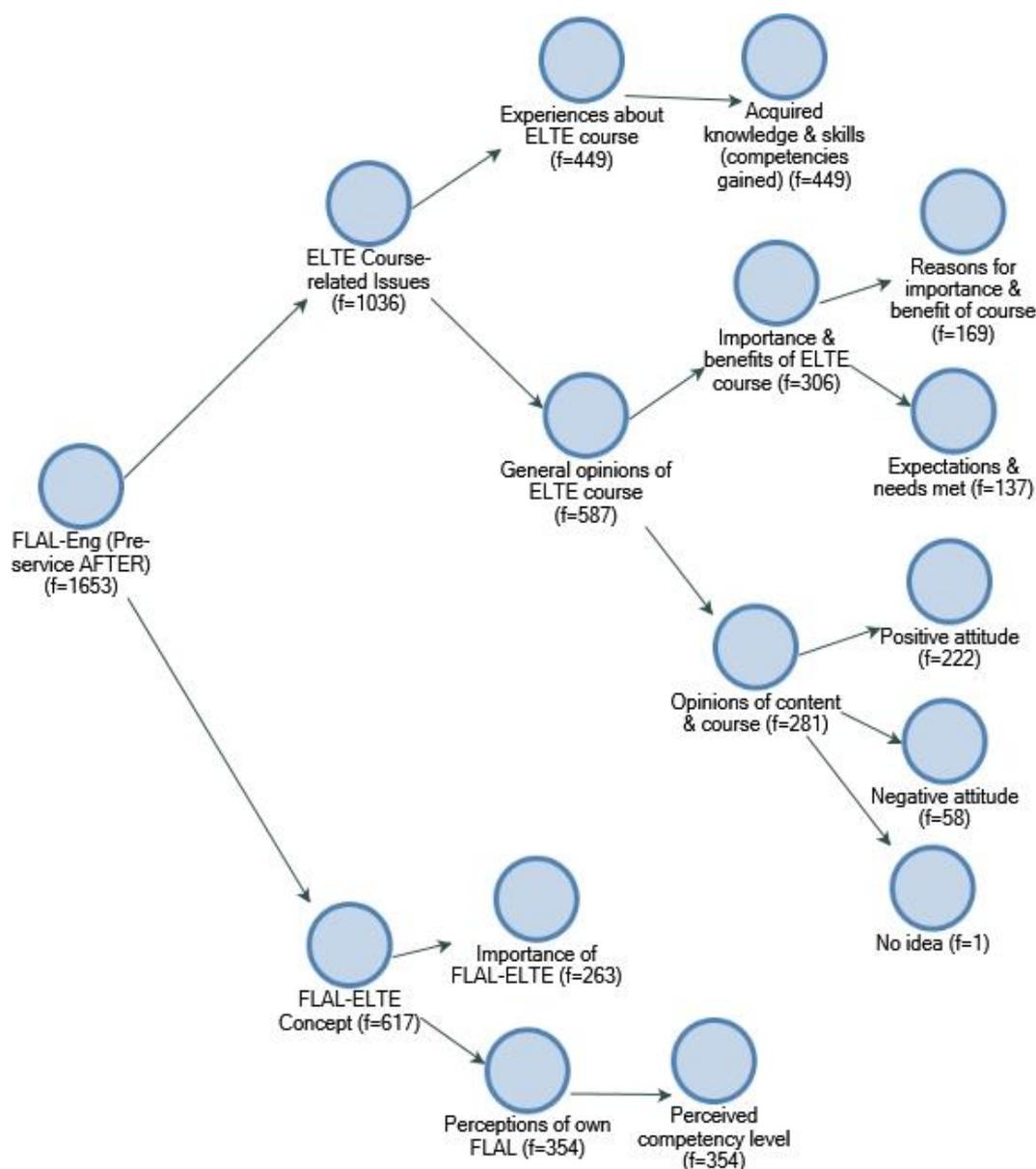

Figure 2. The thematic display of the findings from pre-service ELT teachers (after the course)

For '*FLAL-ELTE Concept*', they underscored the importance of FLAL-ELTE in general (f=263) and mentioned their own perceived competency level (f=354) (see Figure 2). Most participants thought testing and evaluation are very important in teaching process because they are useful to provide feedback about especially what is right and wrong about learning. Therefore, they stated every teacher, no matter what they are candidates, novice or experienced, must acquire language assessment competence *('Importance of FLAL-ELTE')*.

As to '*Perceptions of own FLAL level*', most participants believed they needed to improve themselves more because they felt themselves competent at the moderate level in FLAL after the course (*n*=109). They hoped to make progress in testing especially about practical skills when gaining more experience in their future teaching. For example, one of them indicated:

'I perceived myself being qualified in language assessment at a moderate level. It is because I cannot foresee how effectively I will put theoretical knowledge into practice. I believe I will be better in language assessment via experience.' (After.Moderate.Pre.100)

In contrast to moderate level perception, those who perceived themselves highly competent in FLAL (*n*=33) discussed that though they were still teacher candidates, they assumed to learn necessary things about language testing, and be able to perform tests easily in their teaching life thanks to this course. On the other hand, although there were few participants who perceived themselves to be very inadequate (*n*=4), they complained about their lack of testing experience even after taking that course.

When it comes to '*ELTE Course-related Issues*', the participants expressed their general opinions about the course (*f*=587) and mentioned their experiences throughout the course in the form of what they learned (*f*=449) (see Figure 2).

Concerning '*Experiences about ELTE course: acquired knowledge and skills*', most of them focused on their gained competencies about language assessment in which they discussed what they learned and how they could use it. They reported they learned how to test each language skill, how to design language skills tests by using different test items such as multiple choice items, how table of specifications can be used, and what kind of testing criteria and principles, test types and purposes there are. For instance, one of them exemplified that:

'We received information about the criteria that we have to pay attention while designing language tests. Also, we learned how we can prepare language exams according to the level of our classes as well as our teaching goals. Besides, we studied on the procedures which can be utilized to increase the reliability and validity of language tests.' (After.Moderate.Pre.119)

Also, some stated that their views of testing changed; they said they realized teaching-testing connection and many affecting factors on testing. For example, one indicated that:

'I learned at the end of this course that there is a close relationship between teaching and assessment. I also learned about language testing techniques, evaluation, the features that any measurement tool should possess, and how to prepare test items for language skills such as grammar, vocabulary, reading, writing, and listening.' (After.Moderate.Pre.44)

Considering '*General opinions of ELTE course: its importance and benefits*', nearly all the participants agreed the course responded to their needs and expectations. The course was found beneficial for the career; in addition, they believed they gained necessary competencies about language assessment. Some also mentioned the ELTE course was helpful to provide another teaching competence dimension, namely, language assessment during their teacher education. For instance, one of them discussed:

'This course met my expectations. Throughout our four-year-university life, we took the courses about how to teach English language; I mean, how to organize language teaching activities in the class. But we had not known how to assess those at the end of those courses. From this point, this course filled this gap.' (After.Moderate.Pre.31)

With respect to '*Opinions of content and course*', most of the sample held positive attitudes and found the content sufficient. They indicated the topics were comprehensive, and they learned test types, items, methods, and other stuff while developing an understanding of

the concept. Besides, they highlighted their course instructors were very qualified in testing, and very helpful and explanative throughout the course semester.

Nevertheless, some participants discussed the shortcomings of the course and made suggestions to improve it. They mostly criticized the course was too theoretical and there was no practice; thus, they did not gain practical testing skills during the term. For example, one of them reported:

'The course did not respond to my needs adequately because the whole semester was theoretical. The practical part was given only as an assignment. While the theoretical parts could have been read at home, there was much more emphasis on that side in the lessons. Therefore, since we had no experience in terms of practicing assessment, there would be problems in the future.' (After.Low.Pre.4)

Moreover, some complained about not-updated course content. The coursebook used was Heaton's (2011) '*Writing English Language Tests*', but the participants found the sample testing items old and not suitable to recent teaching contexts. Some participants also mentioned the semester of the course was too late to take because they believed they did not have any chance to practice their testing skills. Even in teaching practice, they did not do any testing practice since they were only evaluated by their teaching skills, not testing skills. Similarly, they complained about the inadequate class hours because the course was delivered only three hours a week only in one semester.

Thereupon, the participants made some recommendations to improve the course. For example, they indicated the practical side could be increased by means of the assignments of preparing and administering language tests. In addition, some suggested there might be more examples related to integrated testing skills since they mentioned they only learned testing each language skill separately. They also stated more checklists can be given to learn how to evaluate and score because there was not much anything about scoring or grading. Besides, some emphasized the course might be coordinated in line with the practicum where students may implement their own language testing tasks. To exemplify;

'The course was sufficient in terms of its theoretical knowledge, but I think it was inadequate in terms of its practical side. I think the course should be integrated into the practicum.' (After.Moderate.Pre.26)

In short, the participants appreciated the importance and benefit of the ELTE course, but they did not feel very qualified in FLAL especially due to the lack of practice and experience. Therefore, they drew attention to some missing points, and accordingly made certain suggestions to develop the training course more.

4. Discussion

The current study investigated LAL training needs of pre-service ELT teachers through their opinions and experiences about ELTE course. The findings yielded that ELTE course provided a basic training and offered essential theories of language testing and knowledge of preparing specific language-skill-based tests. Most participants found the course beneficial, especially for its being English-domain-focused and complementary to teaching-methods-courses. Thus, there was a positive change after the course because they mentioned they learned a lot and were more aware and critical of ELTE, which is similar to Hilden and Forejdendahl's (2018), and Watanabe's (2011) study results. In terms of their perceived competency level, most of the participants felt competent at the moderate level though they needed to experience assessment, which is contrary to some study findings that revealed

lower levels of LAL (e.g. Howerton, 2016; Karaman & Sahin, 2014). Besides, they indicated their course instructors were qualified in ELTE, which helps them to gain necessary knowledge. This is similar to Jeong's (2013) argument that the background of teacher educators is important in that they must be professionally developed to give content courses. Hence, the ELTE course was found to have a positive effect in changing attitudes and gaining certain knowledge of testing as perceived by the sample. Yet, the participants discussed some shortfalls related to the course and certain LAL training needs were identified.

To begin with the conceptions, pre-service ELT teachers adopted a rather testing notion than assessment perspective because they mostly focused on summative purposes of assessment such as diagnosis and achievement/attainment that can be measured via traditional testing tools like multiple-choice items. Similarly, some studies also found that ELTE was seen only as summative assessment (e.g. Lam, 2014; Viengsang, 2016). Actually, this testing perception can be attributed to two reasons: First, the ELTE course content included more testing topics in line with the coursebook and hence, their perceptions were shaped accordingly. Second, Turkey is an exam-oriented country which employs high-stakes testing through multiple-choice exams; thus, the participants were more familiar with that due to their previous experiences. This was also highlighted by Lam (2014) that country's tendency in assessment affects perceptions of teacher candidates. In addition, regarding LAL dimensions, pre-service ELT teachers emphasized knowledge/what component and skills/practices/how component more than principles (proper/ethical use) and contexts (origins, reasons/why component, impacts, teaching settings) compared to Davies's (2008), Fulcher's (2012), and Inbar-Lourie's (2008) frameworks. In other words, they did not underline much some LAL characteristics such as ethics, fairness, knowledge of social/political concepts, rationales behind assessment choices, backwash effect and contextual or local needs. Therefore, there is a need to broaden their perspectives towards more formative assessment viewpoint to keep up with the recent understandings of assessment for learning; that is, assessment can be used to foster language learning apart from just measuring language proficiency, especially required at the classroom level (Fulcher, 2012; Scarino, 2013). There is also a need to concentrate on other LAL components to make them comprehend the elements of being assessment literate for better education.

Considering the topics of the ELTE content, the statements of the participants indicated the syllabus contained testing approaches and methods, testing criteria (reliability, validity, etc.), and the design of language-skill-tests by writing appropriate test items like cloze tests. This was most probably due to the testing coursebook (Heaton, 2011). This is similar to the findings of Jin's (2010) and Hatipoglu's (2015) study that such courses consisted of only testing matters. Hence, other topics such as alternative, performance-based, computer-based assessment, statistics, the comparisons of teacher-made and ready-made tests or classroom-based and large scale/standardized tests, contextual issues, and ethical concerns were not touch upon much in the training. Specifically, the participants stressed they learned how to test each skill separately and did not know how to measure overall language ability. They also stated they did not learn anything about scoring, grading, and evaluating the results to give feedback, which is similar to Watanabe's (2011) study that pre-service teachers had difficulties in communicating with assessment results. However, one of the characteristics of being LAL is the ability to analyze and interpret assessment results to improve and reinforce learning (Berry & O'Sullivan, 2016; Taylor, 2009). For this reason, there is a need to revise the syllabus of the course and to add some other materials/resources for a better LAL training.

Moreover, the participants complained about the inadequate class hours and being too late to take such training in the last semester as in Hatipoglu's (2015) study because the content

was too demanding, and there was not much time to achieve each objective of the course throughout the semester. Thus, it can be inferred all the components of LAL cannot be focused in a single course and there is a need to redesign course hours and given semester time.

Furthermore, although the participants reported the content was sufficient and their expectations were met, they agreed the course was too theoretical and lacked practice as in some studies (e.g. Jin, 2010; Komur, 2018; Lam, 2014). For example, they criticized though they learned how to design a grammar multiple-choice test item, they did not perform such knowledge in the course. Apart from the course, they did not even practice their testing knowledge in the practicum as also reported by Sariyildiz (2018) and Viengsang (2016) because they were only evaluated by their teaching in the practicum. In fact, this is against the interconnection between teaching and assessment. Since there was imbalance between theory and practice in the course, practical testing skills were not developed much. Therefore, further training is needed to enable them to experience what they learn. Otherwise, all they learned remained in theory as in the present study.

To sum up, concerning the first research question, the needs of pre-service ELT teachers about LAL training were listed as: deprived of having assessment and evaluation perspective on the conceptual basis, lack of knowledge about formative assessment, ethical issues, contextual and local uses of assessment, and being deficient in putting theory into practice. As for the second research question, ELTE course was found useful to make significant contributions to develop ELT-based-testing knowledge by teacher candidates, especially in terms of raising awareness with regard to language testing principles, types, and items, learning how to design testing tools about language skills and areas, and leading them to feel more competent in testing knowledge. Conversely, they highlighted negative aspects of the course as insufficient hands-on experience in designing and administering tests, and interpreting test results, incapable of forming integrated-skills tests, inadequate class hours, and its being given too late in the program. Since the content included only testing knowledge issues at the theoretical base, it was restricted in covering the following topics: recent assessment means such as formative assessment, computer-based testing, statistical procedures, and alternative assessment such as portfolio, self/peer-assessment, and learner diaries.

Though it is promising to see that ELTE course provided an initial basic training and positive changes in teacher candidates, still there is a need to revise the course and make it more practice-based to make future teachers more competent because a good LAL training should include balanced proportions of what-how-why of language assessment depending on a contextualized instruction (Inbar-Lourie 2008; Taylor 2009).

Therefore, some implications can be shared. Firstly, different meanings and functions of language assessment can be explicitly introduced and exemplified to broaden the perceptions. Secondly, other topics such as integrated testing, formative and alternative assessment might be added to update the content. Likewise, local needs of the educational system and unethical assessment practices may be discussed and exemplified to boost LAL more. To enhance evaluation skills, the preparation of rubrics/checklists or scoring/grading ready-made materials can also be included. Additionally, more examples, exercises and tasks might be studied to make them to connect their assessment knowledge with their skills, and then they can construct own tools and administer them to internalize what they learn. It is because without practice, teacher candidates cannot develop meaningful assessment. However, to feel real experiences, there should be a coordination between ELTE course and teaching practice through which they can put their theoretical assessment knowledge into practice through the

cycle of plan-develop-administer-evaluate language assessment according to their lessons in the practicum.

In view of such implications, a model content can be recommended for a better course-based LAL training for teacher candidates of ELT. In this model, the course can be divided into two semesters as two following courses ELTE-I (7th semester) and ELTE-II (8th semester) to handle the demanding content effectively and allocate more time for both knowledge and practice. It should be noted it would be more useful to place the courses in the last two semesters after taking all teaching methodology courses so that it becomes more meaningful to pre-service teachers to relate them. ELTE-I might cover more theoretical issues, and ELTE-II more practical considerations, which also includes its integration into practicum; that is, while they take in-class instruction, they can practice assessment in practicum.

ELTE-I may contain introduction into language testing, assessment, and evaluation; assessment of learning and assessment for learning; the concept of LAL; historical background, recent developments, and social and contextual issues; principles of language testing (e.g., reliability, validity, washback, ethics); purposes, types and uses of multiple language tests and assessments (e.g., achievement, placement, summative, formative, alternative, high/low-stakes, classroom-based, standardized, computer-based, teacher-made, etc.); commonly used types of test items and exercises on their examples; the ways of performing formative assessment (e.g., self/peer-assessment, observation, portfolio); introduction to evaluation; the ways of analysis and interpretation of scores with sample exercises by using basic descriptive and inferential statistics; the ways of evaluating and monitoring language progress and giving feedback.

ELTE-II might consist of the revision of ELTE-I; test construction stages and the use of table of specifications; summative and formative assessment ways of language areas (e.g., grammar, vocabulary, pronunciation), skills (e.g., reading, listening, speaking, writing), and integrated one by studying on samples, doing exercises and activities; constructing tests and criteria; discussing the practicum experiences and observations; reflecting on the performance of pre-service teachers' tests and assessment ways in teaching practice by means of administering and interpreting them (i.e., the coordination with practicum through the cycle of plan-develop-administer-evaluate language assessment; not just summative but also using formative one).

After all, since teacher education is influential in shaping early beliefs and training ELT specific competencies, which affects future performance of teacher candidates, such revisions in response to the needs of pre-service ELT teachers might be helpful to promote their LAL for a better teaching career.

5. Conclusion

Language teachers need to be assessment literate to respond to the needs of their educational context more effectively because LAL is helpful to undertake efficient assessment methods, which makes both teaching and learning to be developed and motivated for better outcomes (Khadijeh & Amir, 2015). But to be assessment literate, they need training early in learning so that they can construct their assessment knowledge and skills for their career. Thus, teacher education has an important role to equip teachers with LAL competence (DeLuca & Klinger, 2010). In this sense, this study was believed to enlighten certain LAL training needs of pre-service ELT teachers within ELTE course through their own perspectives as a distinctive feature and revealed a need to redesign the course in terms

of content, implementations, time, and practice. Unlike other studies in this field, the current study made an effort to share pedagogical implications that can be applicable regarding LAL training through the model course content by including the coordination with practicum so as to provide assessment practices. Specifically, this model content for the courses as ELTE-I and ELTE-II could lead to a new curriculum design for pre-service teachers' LAL training in teacher education program. This model is also hoped to be implemented in other contexts by making necessary adaptations according to their own educational policies. In this respect, this study showed a genuine attempt to present a new model for a course-based LAL training by starting from a specific context (Turkish-ELT) to guide the development of such courses in other countries.

As suggestions for further studies, since this study was a sample of only one stakeholder, the perspectives of different stakeholders can be considered to elaborate more on the needs of LAL training; for instance, comparisons between pre-service and in-service teachers regarding their perceptions and practices of LAL can be made to improve the quality of training and provide more guidelines for teacher education.

References

Bachman, L. F. (2005). *Statistical analyses for language assessment*. CUP.

Berger, A. (2012). Creating language-assessment literacy: A model for teacher education. In J. Hüttner, B. Mehlmauer-Larcher, S. Reichl & B. Schiftner (Eds.), *Theory and practice in EFL teacher education: Bridging the gap* (pp. 57-82). Short Run Press.

Berry, V., & O'Sullivan, B. (2016). *Assessment literacy for language teachers*.

http://tea.iatefl.org/wp-content/uploads/2015/10/Vivien-Berry-and-Barry-OSullivan.pdf

Bohn, H., & Tsagari, D. (2021). Teacher educators' conceptions of language assessment literacy in Norway. *Journal of Language Teaching and Research 12*(2), 222-233. http://dx.doi.org/10.17507/jltr.1202.02

Chen, P. P. (2005). Teacher candidates' literacy in assessment. *Academic Exchange Quarterly, 9*(3), 62-67. https://www.thefreelibrary.com/Teacher+candidates%27+literacy+in+assessment.-a0138703662

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson Education.

Csépes, I. (2014). Language assessment literacy in English teacher training programmes in Hungary. In J. Hovarth & P. Medgyes, (Eds.), *Studies in Honour of Marianne Nikolov* (pp. 399-411). Lingua Franca Csoport.

Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing, 25*(3), 327-347. https://doi.org/10.1177%2F0265532208090156

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice, 17*(4), 419-438. https://doi.org/10.1080/0969594X.2010.516643

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

Gan, L., & Lam, R. (2020). Understanding university English instructors' assessment training needs in the Chinese context. *Language Testing in Asia 10*(11), 1-18. https://doi.org/10.1186/s40468-020-00109-y

Gay, L. R., Mills, G. E., & Airasian, P. W. (2012). *Educational research: Competencies for analysis and applications* (10th ed.). Pearson Education.

Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile: Issues in Teachers' Professional Development, 20*(1), 179-195. https://doi.org/10.15446/profile.v20n1.62089

Giraldo, F., & Murcia, D. (2018). Language assessment literacy for pre-service teachers: Course expectations from different stakeholders. *Gist Education and Learning Research Journal, 16*, 56-77.

Hatipoglu, C. (2010). Summative evaluation of an English language testing and evaluation course for future English language teachers in Turkey. *English Language Teacher Education and Development* (*ELTED), 13,* 40-51.

Hatipoglu, C. (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.

Heaton, J. B. (2011). *Writing English language tests* (new ed.). Longman.

Herrera, L., & Macias, D. (2015). A call for language assessment literacy in the education and development of teachers of English as a foreign language. *Colombian Applied Linguistics Journal, 17*(2), 302-312. https://doi.org/10.14483/udistrital.jour.calj.2015.2.a09

Hidri, S. (2018). Introduction: State of the art of assessing second language abilities. In S. Hidri (Ed.), *Revisiting the assessment of second language abilities: From theory to practice* (pp. 1-19). Springer International Publishing.

Hilden, R., & Frojdendahl, B. (2018). The dawn of assessment literacy – exploring the conceptions of Finnish student teachers in foreign languages. *Apples – Journal of Applied Language Studies, 12*(1), 1–24. https://doi.org/10.17011/apples/urn.201802201542

Hill, K. (2017). Understanding classroom-based assessment practices: A precondition for teacher assessment literacy. *Papers in Language Testing and Assessment, 6*(1), 1-17.

Howerton, A. M. (2016). *Elephant on a stepladder: an exploration of pre-service English teacher assessment literacy* [Doctoral dissertation, Northern Illinois University]. Northern Illinois University Institutional Repository. https://commons.lib.niu.edu/handle/10843/20902

Huang, J., & He, Z. (2016). Exploring assessment literacy. *Higher Education of Social Science, 11*(2), 18-27. http://dx.doi.org/10.3968/8727

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing, 25*(3), 385-402. https://doi.org/10.1177%2F0265532208090158

Inbar-Lourie, O. (2013). Language assessment literacy. In C. A. Chapella (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1-9). Blackwell Publishing. https://doi.org/10.1002/9781405198431.wbeal0605

Inbar-Lourie, O. (2017). Language assessment literacy. In E. Shohamy, L. Or, & S. May (Eds.), *Language Testing and Assessment* (pp. 257-270). Springer International Publishing. https://doi.org/10.1007/978-3-319-02261-1_19

Jeong, H. (2013). Defining assessment literacy: Is it different for language testers and non-language testers? *Language Testing, 30*(3), 345-362. https://doi.org/10.1177%2F0265532213480334

Jin, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing, 27*(4), 555-584. https://doi.org/10.1177%2F0265532209351431

Karaman, P., & Sahin, C. (2014). Ogretmen adaylarinin olcme degerlendirme okuryazarliklarinin belirlenmesi [Investigating the assessment literacy of teacher candidates]. *Ahi Evran Universitesi Kirsehir Egitim Fakultesi Dergisi [Ahi Evran University Kirsehir Education Faculty Journal], 15*(2), 175-189.

Khadijeh, B., & Amir, R. (2015). Importance of teachers' assessment literacy. *International Journal of English Language Education, 3*(1), 139-146. https://doi.org/10.5296/ijele.v3i1.6887

Komur, S. (2018). Preservice English teachers' assessment awareness: Level of readiness for classroom practice. *The Journal of Language Teaching and Learning, 8*(1), 109-121.

Lam, R. (2014). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing, 32*(2), 169-197. https://doi.org/10.1177%2F0265532214554321

Malone, M. E. (2008). Training in language assessment. In E. Shohamy & N. H. Hornberger (Eds.), *The Encyclopedia of Language and Education* (2nd ed., Vol. 7) (pp. 225-239). Blackwell Publishing.

Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi-Ankara Universitesi TOMER [Language Journal-Ankara University TOMER], 168*(1), 43-60.

Miles, M. B., & Huberman, A. M. (1994). *An expanded sourcebook: Qualitative data analysis* (2nd ed.). Sage Publications.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage Publications.

Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing, 30*(3), 381-402. https://doi.org/10.1177%2F0265532213480337

Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing, 21*(3), 249-258. https://doi.org/10.1191%2F0265532204lt283ed

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

Royse, D., Staton-Tindall, M., Badger, K., & Webster, J. M. (2009). *Needs assessment*. OUP.

Sariyildiz, G. (2018). *A study into language assessment literacy of pre-service English as a foreign language teachers in Turkish context* [Unpublished master's thesis]. Hacettepe University.

Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing, 30*(3), 309-327. https://doi.org/10.1177%2F0265532213480128

Stabler-Havener, M. L. (2018). Defining, conceptualizing, problematizing, and assessing language teacher assessment literacy. *Teachers College, Columbia University Working Papers in Applied Linguistics & TESOL, 18*(1), 1-22.

Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics, 29*, 21-36. https://doi.org/10.1017/S0267190509090035

Tsagari, D., & Vogt, T. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment, 6*(1), 41-63.

Turk, M. (2018). *Language assessment training level and perceived training needs of English language instructors: A mixed methods study* [Unpublished master's thesis]. Bahcesehir University.

Ukrayinska, O. (2018). Developing student teachers' classroom assessment literacy: The Ukrainian context. In S. Hidri (Ed.), *Revisiting the assessment of second language abilities: From theory to practice* (pp. 351-371). Springer International Publishing.

Viengsang, R. (2016). Exploring pre-service English teachers' language assessment literacy. *Modern Journal of Language Teaching Methods, 6*(5), 432-442.

Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly, 11*(4), 374-402. https://doi.org/10.1080/15434303.2014.960046

Wach, A. (2012). Classroom-based language efficiency assessment: A challenge for EFL teachers. *Glottodidactica, 39*(1), 81-92.

Watanabe, Y. (2011). Teaching a course in assessment literacy to test takers: Its rationale, procedure, content and effectiveness. *Cambridge ESOL: Research Notes, 46,* 29-34.

Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education, 58,* 149-162. https://doi.org/10.1016/j.tate.2016.05.010

Yan, X., Zhang, C., & Fan, J. J. (2018). "Assessment knowledge is important but …": How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System, 74,* 158-168. https://doi.org/10.1016/j.system.2018.03.003