

EXAMINING TURKISH EFL TEACHERS' RECEIVED TRAINING AND PROFESSIONAL DEVELOPMENT NEEDS IN LANGUAGE ASSESSMENT

ANÁLISE DA FORMAÇÃO E NECESSIDADES DE DESENVOLVIMENTO PROFISSIONAL DE PROFESSORES TURCOS DE INGLÊS COMO LÍNGUA ESTRANGEIRA NO ÂMBITO DA AVALIAÇÃO LINGÜÍSTICA

ANÁLISIS DE LA FORMACIÓN Y LAS NECESIDADES DE DESARROLLO PROFESIONAL DE LOS MAESTROS TURCOS DE INGLÉS COMO LENGUA EXTRANJERA EN EL ÁMBITO DE LA EVALUACIÓN LINGÜÍSTICA



Merve BAYRAM¹

e-mail: mervebyrm1996@gmail.com



İsmail Fırat ALTAY²

e-mail: ifaltay@hacettepe.edu.tr

How to reference this paper:

Bayram, M., & Altay, I. F. (2026). Examining Turkish EFL teachers' received training and professional development needs in language assessment. *Revista on line de Política e Gestão Educacional*, 30(esp2), e026087. <https://doi.org/10.22633/rpge.v30iesp2.21397>



| Submitted: 14/03/2026

| Revisions required: 22/05/2026

| Approved: 16/06/2026

| Published: 01/08/2026

Editor: Prof. Dr. Sebastião de Souza Lemes

Deputy Executive Editor: Prof. Dr. José Anderson Santos Cruz

¹ Hacettepe University, Ankara – Turkey. Graduate School of Educational Sciences.

² Hacettepe University, Ankara – Turkey. Faculty of Education, ELT Department.

RPGE – Revista on line de Política e Gestão Educacional, Araraquara, v. 30, n. esp. 2, e026087, 2026.

DOI: 10.22633/rpge.v30iesp2.21397

ABSTRACT: This study examines the training and professional development needs of Turkish teachers of English as a foreign language (EFL) in the field of language testing and assessment (LTA). The participants were 100 practicing EFL teachers from different educational levels who completed an online questionnaire on Language Testing and Assessment. The instrument assessed both the training received and professional development needs across four dimensions of language assessment. The data were analyzed quantitatively using descriptive statistics. The results indicate that, overall, the teachers considered their prior training sufficient across the different dimensions of LTA, although variations were observed among the domains. The highest levels of training were reported in relation to test purposes, whereas the lowest levels concerned knowledge of testing and assessment. Despite considering their preparation sufficient, the participants reported basic professional development needs across all assessment dimensions, particularly regarding classroom-based assessment and the more technical and specialized aspects of language assessment.

KEYWORDS: Language assessment literacy. Language testing and assessment. EFL teachers. Professional development. Teacher education.

RESUMO: O presente estudo examina a formação e as necessidades de desenvolvimento profissional de professores turcos de inglês como língua estrangeira (EFL) na área de testes e avaliação de línguas (LTA). Participaram 100 professores de EFL em atividade, de diferentes níveis de ensino, que responderam a um questionário on-line sobre Testes e Avaliação de Línguas. O instrumento avaliou a formação recebida e as necessidades de desenvolvimento profissional em quatro dimensões da avaliação de línguas. Os dados foram analisados quantitativamente por meio de estatísticas descritivas. Os resultados indicam que, de modo geral, os professores consideraram suficiente sua formação prévia nas diferentes dimensões da LTA, embora tenham sido observadas variações entre os domínios. Os níveis mais elevados de treinamento foram relatados quanto aos objetivos dos testes, enquanto os mais baixos se referiram ao conhecimento sobre testes e avaliação. Apesar da preparação considerada suficiente, os participantes apontaram necessidades básicas de desenvolvimento profissional em todas as dimensões, especialmente na avaliação em sala de aula e nos aspectos técnicos e especializados da avaliação de línguas.

PALAVRAS-CHAVE: Letramento em avaliação de línguas. Testes e avaliação de línguas. Professores de Inglês como Língua Estrangeira. Desenvolvimento profissional. Formação de professores.

RESUMEN: El presente estudio examina la formación y las necesidades de desarrollo profesional de profesores turcos de inglés como lengua extranjera (EFL) en el ámbito de las pruebas y la evaluación de lenguas (LTA). Participaron 100 profesores de EFL en ejercicio, pertenecientes a diferentes niveles educativos, quienes respondieron a un cuestionario en línea sobre Pruebas y Evaluación de Lenguas. El instrumento evaluó tanto la formación recibida como las necesidades de desarrollo profesional en cuatro dimensiones de la evaluación de lenguas. Los datos se analizaron cuantitativamente mediante estadística descriptiva. Los resultados indican que, en general, los profesores consideraron suficiente su formación previa en las diferentes dimensiones de la LTA, aunque se observaron variaciones entre los distintos ámbitos. Los niveles más altos de formación se registraron en relación con los objetivos de las pruebas, mientras que los más bajos correspondieron al conocimiento sobre pruebas y evaluación. A pesar de considerar suficiente su preparación, los participantes señalaron necesidades básicas de desarrollo profesional en todas las dimensiones de la evaluación,

especialmente en la evaluación en el aula y en los aspectos técnicos y especializados de la evaluación de lenguas.

PALABRAS CLAVE: Alfabetización en evaluación lingüística. Pruebas y evaluación lingüística. Docentes de inglés como lengua extranjera. Desarrollo profesional. Formación docente.

INTRODUCTION

Language testing and assessment play a key role in the language learning process by shaping teachers' decisions, guiding the practices in the classroom, and influencing students' language achievement. In addition to evaluating student performance, assessment provides teachers with valuable information that enables them to improve their teaching strategies and enhance learning (Stiggins, 2008). The correlation between language learning and assessment is therefore complementary, as assessment processes have an impact on teaching methods, whereas language teaching approaches also affect assessment practices (Cheng, 1997; Clapham, 2000; Spolsky, 1995). Within this dynamic relationship, effective language teaching requires teachers to possess the necessary knowledge and skills to create appropriate assessment practices and interpret students' performance (Stiggins, 2014).

This requirement emphasizes the concept of "Assessment Literacy", which refers to teachers' capacity to recognize the goals of assessment, make appropriate decisions, evaluate student performance successfully and prevent potential issues in assessment procedures (Stiggins, 1995). Originally coined by Stiggins (1995), assessment literacy involves not only conceptual knowledge about assessment principles, but also competencies for appropriately and effectively implementing these principles in classroom settings (Siegel & Wissehr, 2011). From a pedagogical perspective, Borg (2015) further framed assessment literacy as a fundamental component of teacher cognition.

In language education, the broader term of assessment literacy has been further specified as Language Assessment Literacy (LAL), which emphasizes the field-specific understanding, competencies, and principles necessary for language assessment. It is described as "the ability to design, develop, and critically evaluate tests and other assessment procedures, as well as the ability to monitor, evaluate, grade, and score assessments on the basis of theoretical knowledge" (Vogt & Tsagari, 2014, p. 377).

In this regard, LAL is recognized as an essential component of teachers' professional development (Popham, 2009) and as a characteristic of effective teachers (Mertler & Campbell, 2005). With the increasing emphasis on formative and learning-centered assessment and socio-cultural perspectives in education, LAL has emerged to address the assessment challenges faced by language teachers and to highlight its importance in foreign language education (Inbar-Lourie, 2017). Consequently, the significance of LAL as a required professional competence has expanded in recent years due to language teachers' increasing identity as assessors (Fulcher, 2012).

Despite the increasing awareness of LAL as a crucial element of successful language education, empirical research on teachers' competencies has yielded inconsistent findings, and the existing research, particularly in the Turkish context, remains relatively inadequate. Some studies indicate that teachers demonstrate overall insufficient or uneven levels of LAL across different domains of language assessment and report a need for further professional development (Ölmezer-Öztürk & Aydın, 2019; Şahin, 2019; Tsagari & Vogt, 2017; Vogt & Tsagari, 2014).

By contrast, other studies reveal that although teachers are equipped with sufficient theoretical understanding, they require assistance in applying this knowledge in the classroom (Öz & Atay, 2017; Shim, 2009). Beyond these inconsistent findings, the rationale for the present research arises from the limited scope of the existing literature. Although most LAL research in Turkey examines pre-service teachers or instructors working in higher education, relatively limited research has been conducted on in-service teachers in primary, secondary, and high schools who actively participate in assessment activities.

Therefore, further investigation is required to examine the current levels of LAL and the professional development needs of EFL teachers, in response to the recent changes in education and assessment policies. To address this gap, the present study aims to examine EFL teachers' received training in language assessment and related professional development needs. Based on these objectives, the following research questions were formulated:

RQ1: What are the levels of language assessment literacy among in-service EFL teachers?

RQ2: What are in-service EFL teachers' perceived training needs across different dimensions of language assessment?

LANGUAGE ASSESSMENT LITERACY

Language assessment tracks students' progress and shapes educational decisions and outcomes; therefore, it is a vital component of language teaching and learning processes. In addition to possessing strong pedagogical skills, language teachers, as "language assessors" (Fulcher, 2012), must understand and successfully implement the assessment techniques and concepts that significantly support students' development and guide decision-making regarding instruction. In this sense, Language Assessment Literacy (LAL) has emerged as a distinct and fundamental competency that is increasingly viewed as an indispensable component of

language teachers' professional development (Fulcher, 2012; Kremmel & Harding, 2020; Popham, 2009).

In its broadest sense, LAL refers to the concepts, skills, and knowledge that language teachers need to plan, conduct, interpret, and analyze assessment processes appropriately and effectively. The significance of LAL in modern language education has been further reinforced by the increasing focus on formative and learner-centered assessment, as well as by sociocultural considerations (Inbar-Lourie, 2017). As teachers typically play a key role in assessment procedures, their assessment literacy has a direct impact on the effectiveness of classroom practices and student achievement.

Language Assessment Literacy, which has its roots in Stiggins's (1991) introduction, is consistent with AL and may even be linked to it (Taylor, 2013). While assessment literacy provides a broad framework, LAL further expands this notion by integrating the distinctive features of language as a means of communication and a subject of education. Although LAL and assessment literacy share certain elements and standard features, LAL is distinct from assessment literacy regarding the nature of language use, pedagogy, and knowledge (Brookhart, 2001; Giraldo, 2018).

In this way, LAL encompasses fundamental assessment principles such as measurement, validity, and reliability, ensuring not only a thorough understanding of these concepts but also their application through methods that are pedagogically vital in language education. Rather than being a static body of information, the existing literature generally describes LAL as a multifaceted and dynamic structure. Furthermore, LAL is now recognized not merely as a set of specific abilities, but as a concept framed by teachers' professional settings, experiences, attitudes, and values (Scarino, 2013). In this sense, LAL can be conceptualized as a competence embedded in a social context and shaped by not only teachers' classroom practices but also their beliefs and biases (Giraldo, 2018).

Language assessment literacy encompasses not only technical and educational aspects but also the critical and ethical elements of assessment procedures (Inbar-Lourie, 2008; Stabler-Havener, 2018; Vogt & Tsagari, 2014). To describe these components of LAL, several conceptual frameworks have been developed in the literature (Brindley, 2001; Davies, 2008; Fulcher, 2012; Inbar-Lourie, 2008). While some scholars characterize LAL as multi-component models (Brindley, 2001; Davies, 2008; Fulcher, 2012; Inbar-Lourie, 2008), others view LAL progression as a continuing procedure (Pill & Harding, 2013).

Three fundamental elements—knowledge, skills, and principles—are shared by most frameworks, although there are variations in their scope and technical terms. In addition, some frameworks further expanded the conceptualization of LAL by integrating the wider political, social, and organizational settings in which the assessment was conducted. Similarly, Fulcher's (2012) framework emphasizes assessment techniques, assessment principles, and context, whereas developmental perspectives conceptualize LAL as a competency that changes over time as a result of experience, professional involvement, and contextual factors (Pill & Harding, 2013; Taylor, 2013; Yan & Fan, 2021). Taken together, LAL can be viewed as a complicated, dynamic construct that integrates theoretical understanding, practical implementation, ethical considerations, and contextual awareness.

PREVIOUS STUDIES ON LANGUAGE ASSESSMENT LITERACY

International Context

As the role of testing and assessment in language education has gained increasing recognition, LAL has become a growing area of interest in international research over the past two decades (Fulcher, 2012; Jeong, 2013). Language Assessment Literacy has been investigated across diverse stakeholder groups, including pre-service and in-service teachers as well as other professionals participating in language assessment procedures, in many different countries and across various educational contexts. Numerous research topics related to LAL have been covered in international studies, involving the efficacy of testing courses in teacher education (Bailey & Brown, 1996; Brown & Bailey, 2008), knowledge of language assessment (Xu & Brown, 2017), teachers' beliefs and practices regarding assessment (Rogers et al., 2007), perceived LAL levels and training requirements (Hasselgreen et al., 2004; Vogt & Tsagari, 2014), the development of LAL (Levi & Inbar-Lourie, 2020; Tian et al., 2021; Yan & Fan, 2021) and factors influencing LAL development (Crusan et al., 2016; Sultana, 2019).

Among these topics, international empirical studies generally aim to investigate teachers' LAL profiles and identify gaps in assessment training. The results of these studies demonstrate that while teachers perceive themselves as competent in some areas of language assessment, their overall LAL levels are frequently identified as moderate or limited, along with an ongoing demand for additional education. These findings indicate that LAL is shaped by a variety of dimensions and conditions, highlighting its multi-faceted and context-sensitive structure.

Both comprehensive and context-specific research generally reveal that teachers identify significant limitations in more specialized and conceptual areas of language assessment, while they perform adequately in regular classroom-based assessment. In comprehensive European research conducted with different stakeholder categories in 37 countries, Hasselgreen et al. (2004) found that language teachers reported considerable training needs across different dimensions of language assessment, particularly in areas such as test development, interpretation of results, statistics, and alternative assessment, while expressing greater confidence in routine classroom assessment practices.

Drawing on this study, Vogt and Tsagari (2014) investigated teachers' LAL across the same dimensions in seven European nations and similarly found substantial needs for further training, especially in classroom-focused assessment, alternative assessment, statistics, validity, and reliability. Further qualitative findings suggested that insufficient formal assessment education often leads teachers to rely on traditional testing methods and personal experience in assessment practices (Tsagari & Vogt, 2017).

Additional information from research carried out in various educational settings confirms these results and highlights the multifaceted structure of LAL. For example, Fulcher (2012) conducted a needs analysis with 278 English teachers from various parts of the world and found that, due to the lack of professional training in language assessment, teachers frequently rely on standard books, pre-made sources, and past experience, while also demonstrating limited awareness of key assessment concepts such as quality criteria and statistics. Research carried out in the Chinese setting reveals similar tendencies.

Xu and Brown (2017) found that English language teachers in China showed basic levels of assessment literacy in core areas, whereas Zhang et al. (2026) reported that language teachers possessed greater expertise in classroom assessment processes but lacked knowledge regarding conceptual and technical aspects of assessment, resulting in a significant need for additional professional development. Similar findings have also been reported in other educational contexts. Studies conducted in Ukraine, Hong Kong, and Bangladesh found that teachers demonstrated greater competence in routine classroom assessment practices but experienced difficulties in areas such as validity, reliability, alternative assessment, statistical analysis, and assessment design, often relying on practical experience and sample examinations rather than formal assessment training (Kvasova & Kavytska, 2014; Sultana, 2019; Xie & Tan, 2019).

Overall, these studies suggest that teachers' LAL levels do not develop equally across different dimensions, and although specific needs vary depending on the context, there is a clear need for further professional development in both theoretical and practical facets of assessment.

Turkish Context

Research on LAL in the Turkish context remains relatively limited, despite the vital role that assessment literacy plays in both in-class assessment processes and teacher professionalism. Several research topics related to LAL have been addressed in previous studies, including language assessment training in teacher education (e.g., Hatipoğlu, 2010; Şahin, 2019), teachers' knowledge of language assessment (e.g., Ölmezer-Öztürk & Aydın, 2019), beliefs and practices regarding assessment (e.g., Han & Kaya, 2014; Öz & Atay, 2017; Yastıbaş & Takkaç, 2018), and teachers' perceived levels of LAL and professional development requirements (e.g., Mede & Atay, 2017; Sarıyıldız-Canlı, 2023).

More recent studies have also examined technology-facilitated LAL (Bulut, 2024), online training programs (Bağışlayıcı, 2023), and its relationship with reflective teaching and self-efficacy beliefs (Arıcı, 2022; Çağlayan, 2022). Overall, previous studies have primarily focused on teacher candidates (Sarıyıldız, 2018; Şükür, 2023) and instructors in preparatory programs (Kaya, 2020; Yılmaz, 2021). However, there remains a lack of research on the LALs of in-service EFL teachers and their perceived needs for further education, which provides a clear rationale for the present study.

Several researchers have utilized internationally validated quantitative and qualitative instruments to determine teachers' LAL competency in Turkish educational settings. To illustrate, Mede and Atay (2017) examined the LAL of EFL instructors in preparatory schools using qualitative data and a questionnaire adapted from Vogt and Tsagari (2014). The findings revealed that instructors' levels of LAL were low and also indicated ongoing professional development needs, particularly in relation to classroom assessments and language assessment concepts. In a similar vein, Ballıdağ and İnan-Karagül (2021) used the same questionnaire in mixed-method research with English language teachers at the primary and secondary education.

According to their findings, EFL teachers' received training in language assessment remained inadequate across all domains, indicating only a basic previous education and a continuing demand for additional professional training. On the other hand, they addressed challenges in more specialized areas of assessment, such as alternative assessment, cultural dimensions of assessment, statistical expertise, and issues of reliability and validity. Interviews

with teachers also highlighted inadequate training opportunities with limited practical components and difficulties transferring theoretical knowledge into classroom practice.

A research area more closely related to the topic has focused on the levels of LAL perceived by both pre-service and in-service teachers, as well as their professional development needs, using a validated and widely accepted instrument. Sarıyıldız (2018), for instance, created and used a revised version of the Teacher's Questionnaire (Vogt & Tsagari, 2014) to analyze Turkish pre-service EFL teachers. According to the findings, participants believed their previous education was inadequate in three domains of language assessment, although they expressed comparatively greater strength in specific areas and reported limited opportunities to apply assessment knowledge in practice during teacher training.

Using the same questionnaire and a mixed-methods approach, Sarıyıldız-Canlı (2023) conducted a further investigation with in-service language teachers to determine which areas they had received training in language assessment and which they believed needed additional instruction. The findings revealed that teachers generally did not possess an adequate level of preparation across all LAL domains and reported continuing needs for further professional development. The qualitative findings further suggested that insufficient practical training and difficulties applying theoretical knowledge in classroom contexts were among the main factors limiting teachers' assessment literacy.

METHODOLOGY

Research Design

This study employs a quantitative, cross-sectional, survey design to investigate the LAL levels and training needs of Turkish in-service EFL teachers. This design allows for the investigation of teachers' perceptions at a specific point in time across a relatively large sample. Regarding the aims and research questions, a descriptive function was also employed to investigate EFL teachers' received and needed training in language assessment.

Participants

The participants consisted of 100 in-service Turkish EFL teachers working at various instructional levels across different parts of Turkey. A non-probability sampling strategy integrating both convenience and snowball sampling was used (Dörnyei & Csizér, 2012). Regarding teaching experience, 32% of teachers had 1–5 years, 41% had 6–10 years, 20% had

11–15 years, and 7% had more than 15 years of experience. In terms of teaching level, 25% worked at the primary school level, 39% at the secondary school level, and 36% at the high school level. Regarding educational background, 82% of participants held a bachelor's degree and 18% held a master's degree.

Data Collection Instrument and Procedure

The data were collected using the Language Testing and Assessment (LTA) Questionnaire adapted by Sarıyıldız (2018). The instrument was originally developed by Hasselgreen et al. (2004) and later revised by Vogt and Tsagari (2014). The first part collected participants' demographic information, whereas the second part comprised items measuring teachers' received training and perceived training needs across four dimensions of language assessment: Classroom-focused LTA, Knowledge of Testing and Assessment, Purposes of Testing, and Content and Concepts of LTA. Responses for received training ranged from None to Advanced, while perceived training needs were categorized as None, Basic, Intermediate, or Advanced.

Prior to data collection, permission to use the instrument was received from the responsible author and ethical approval was granted by the Hacettepe University Institute of Educational Sciences Research Ethics Committee. Participation was voluntary, and informed consent was obtained from all participants. Following a pilot administration, the final version of the questionnaire was created using Google Forms and administered online. The questionnaire demonstrated high internal consistency ($\alpha = .970$).

Data Analysis

For quantitative analysis, teachers' responses were numerically coded on a four-point Likert scale ranging from 1 (none) to 4 (advanced level), and mean scores were calculated to describe the training levels teachers received and needed. Quantitative data were analyzed using SPSS 26. To address the first and second research questions, descriptive statistics were employed to investigate EFL teachers' language assessment literacy levels and professional development needs across different dimensions of language assessment. The findings are presented in the following sections with supporting tables.

RESULTS

Before reporting the descriptive findings, cut-off criteria for interpreting the mean scores were specified to enable consistent classification of the level of training teachers received and required in language assessment. Mean scores were assigned in equal intervals with approximately 0.75-point cut-off ranges. In this regard, values from 1.00 to 1.74 reflected no training or no need, 1.75 to 2.49 indicated basic training or need, 2.50 to 3.24 represented sufficient training or intermediate need, and 3.25 to 4.00 demonstrated advanced training or need. Overall descriptive results indicate that EFL teachers generally rated their received training at a sufficient level ($M = 2.97$, $SD = 0.475$), while they still reported a basic level of training need in language assessment ($M = 2.20$, $SD = 0.889$). More specifically, teachers' received and needed training levels were examined across four dimensions of language assessment, as displayed in Table 1. According to the findings, teachers perceived themselves as relatively more qualified in the Purposes of Testing dimension, which had the highest mean score ($M = 3.09$, $SD = 0.570$), while they reported having received a lower level of training in the Knowledge of LTA domain ($M = 2.85$, $SD = 0.610$), which had the lowest mean score, yet still falls within the sufficient range. On the other hand, teachers indicated that they perceived the highest but basic level of need in Classroom-Focused LTA ($M = 2.31$, $SD = 0.933$). In contrast, Purposes of Testing dimension yielded the lowest perceived need ($M = 2.11$, $SD = 0.972$), reflecting a basic level of need.

Table 1.

Descriptive Statistics of EFL Teachers' Received and Needed Training in Four Dimensions

Areas of LTA	Received Training			Needed Training		
	Mean	SE	SD	Mean	SE	SD
1. Classroom-focused LTA	2.91	0.053	0.527	2.31	0.093	0.933
2. Knowledge of LTA	2.85	0.061	0.610	2.24	0.092	0.924
3. Purposes of Testing	3.09	0.057	0.570	2.11	0.097	0.972
4. Content and Concepts of LTA	3.05	0.060	0.596	2.16	0.099	0.995
Total	2.97	0.048	0.475	2.20	0.089	0.889

Note. The Authors.

EFL Teachers' Received and Needed Training in Classroom-focused LTA

Table 2 displays descriptive results for EFL teachers' received and needed training in Classroom-focused LTA. With respect to received training, the findings revealed that teachers feel sufficiently equipped with classroom-based language assessment in pre-service education

($M = 2.91$). In particular, the highest mean scores were found in areas of Adapting Ready Tests ($M = 3.18$), Giving Feedback ($M = 3.17$), and Scoring/Grading ($M = 3.16$), which were close to the upper limit of the sufficient training level.

Similarly, the areas of Preparing Classroom Tests ($M = 3.11$), Using Continuous Assessment ($M = 3.06$), and Preparing Achievement Tests ($M = 3.04$) also demonstrated relatively strong educational background classroom-based assessment. On the other hand, the lowest mean scores in this dimension were observed in Preparing Proficiency Tests ($M = 2.59$, $SD = 0.900$), Preparing Language Aptitude Tests ($M = 2.52$), and Using English Language Portfolio ($M = 2.29$). Among these, the Use of the English Language Portfolio emerged as the only area where 23 teachers reported no prior training while the others indicated having received training generally at a basic level.

Table 2.

Descriptive Statistics of Received and Needed Training in Classroom-focused LTA

LTA Practices	Received Training			Needed Training		
	Mean	SE	SD	Mean	SE	SD
Preparing Classroom Tests	3.11	.055	.549	2.23	.119	1.18
Preparing Diagnostic Tests	2.76	.073	.726	2.51	.111	1.10
Preparing Achievement Tests	3.04	.067	.665	2.29	.116	1.15
Preparing Proficiency Tests	2.59	.090	.900	2.69	.114	1.14
Preparing Placement Tests	2.68	.090	.898	2.65	.115	1.14
Preparing Progress Tests	2.98	.071	.710	2.45	.115	1.14
Preparing Language Aptitude Tests	2.52	.093	.926	2.65	.120	1.20
Using Ready-made Tests	3.13	.072	.720	1.99	.113	1.13
Adapting Ready-made Tests	3.18	.067	.672	2.13	.118	1.17
Stages of Language Test Construction	2.91	.077	.767	2.27	.113	1.12
Scoring-Grading	3.16	.069	.692	2.08	.112	1.11
Giving Feedback	3.17	.064	.637	2.15	.116	1.15
Interpreting Test Scores	2.88	.082	.820	2.21	.109	1.09
Using Self/ Peer Assessment	2.99	.067	.674	2.17	.104	1.04
Using Informal Non-Test Type of Assessment	2.95	.078	.783	2.15	.108	1.07
Using Continuous Type of Assessment	3.06	.069	.694	2.03	.111	1.10
Using English Language Portfolio	2.29	.095	.946	2.61	.115	1.15

Note. The Authors.

Regarding perceived training needs, teachers reported basic-level needs in most areas in classroom-focused LTA, as evidenced by mean scores ranging from 2.0 to 2.6, while a few items fell into the category of intermediate-level training needs. The perceived training needs were highest in more formal and specified assessment areas, including Preparing Proficiency Tests ($M = 2.69$), Preparing Placement Tests ($M = 2.65$), Preparing Language Aptitude Tests

($M = 2.65$), and Using the English Language Portfolio (ELP) ($M = 2.61$), all of which fall into the intermediate need category.

Furthermore, a significant percentage of participants (around one-third) stated that they required an advanced level professional development in these areas. On the contrary, lower training needs were identified for routine classroom assessment practices, such as Using Ready-made Tests, Adapting Ready-made Tests, Using Continuous Type of Assessment, and Scoring-Grading, with mean scores indicating a basic level of need ($M = 1.99$ – 2.13).

EFL Teachers' Received and Needed Training in Knowledge of Testing and Assessment

Descriptive results regarding the level of training EFL teachers have received and need in the second dimension, Knowledge of Testing and Assessment are presented in Table 3. The findings indicate that participants have received sufficient training, as evidenced by the average scores ranging from 2.54 to 3.09 across all components of this dimension, despite showing some variations within the sufficient range. Teachers, in particular, reported training levels within the sufficient range for fundamental assessment constructs, with relatively higher mean scores observed for Formative/Summative Assessment ($M = 3.09$) and Informal/Formal Assessment ($M = 3.08$).

They also indicated moderately sufficient levels of training in Objective/Subjective Testing ($M = 2.98$), Direct/Indirect Testing ($M = 2.89$), and Norm/Criterion-Referenced Assessment ($M = 2.85$). In contrast, mean scores for Computer-Based Testing ($M = 2.54$) and Alternative Assessment ($M = 2.62$) were lower and fell within the sufficient category. These findings were further supported by frequency statistics, indicating that a significant number of teachers responded having received "a little" or "no" training in these areas.

Table 3.

Descriptive Statistics of Received and Needed Training in Testing and Assessment Knowledge

LTA Practices	Received Training			Needed Training		
	Mean	SE	SD	Mean	SE	SD
Informal/ Formal Assessment	3.08	.068	.677	1.98	.108	1.08
Formative/ Summative Assessment	3.09	.070	.698	2.02	.109	1.09
Norm/Criterion-referenced Assessment	2.85	.076	.757	2.20	.107	1.07
Discrete Point/ Integrative Testing	2.75	.078	.783	2.32	.104	1.04
Direct/ Indirect Testing	2.89	.080	.803	2.15	.106	1.05
Objective/ Subjective Testing	2.98	.079	.791	2.13	.104	1.04
Approaches To Language Testing	2.82	.078	.783	2.43	.106	1.05
Alternative Assessment	2.62	.083	.826	2.39	.107	1.07

Computer Based Testing	2.54	.088	.881	2.50	.110	1.09
------------------------	------	------	------	------	------	------

Note. The Authors.

In relation to professional development needs, the areas with the lowest perceived training needs in this dimension were traditional assessment techniques, as shown by the lowest mean scores. In this regard, participants indicated basic level needs for traditional assessment practices such as Formative/Summative Assessment ($M = 2.02$) and Informal/Formal Assessment ($M = 1.98$), as confirmed by high frequency of “none” responses showing no perceived training need. Notably, Computer-Based Testing ($M = 2.50$) was the only area that surpassed the basic need threshold and showed an intermediate level of training need, whereas other items in this dimension remain within the basic need range. Among these, Approaches to Language Testing ($M = 2.43$) and Alternative Assessment ($M = 2.39$) approaches emerged as the areas with the highest demand for professional development among teachers.

EFL Teachers' Received and Needed Training in Purposes of Testing

Table 4 summarizes the degree of training teachers have received and need regarding the third dimension of language assessment, namely the Purposes of Testing. The results revealed that teachers perceived themselves as having the strongest expertise in the areas of Identifying What Needs to Be Learned/Taught ($M = 3.32$) and Identifying What Has Been Learned ($M = 3.28$). These two purposes were classified as “advanced” level training since they were above the upper limit of the “sufficient” range. The frequency results further explained this finding, which illustrate that the majority of participants had “sufficient” ($f = 56/54$) or “advanced” ($f = 38/37$) preparation in these purposes, and that none of the teachers indicated receiving no training.

Giving Grades ($M = 3.19$) and Testing Competence in a Language ($M = 3.15$) were additional testing goals that were around the higher end of the “sufficient” range. Additionally, moderately adequate instruction is indicated by the scores on Measuring General Ability to Learn an L2 ($M = 3.01$) and Placing Students onto Programs ($M = 2.89$). On the other hand, Awarding Final Certificates ($M = 2.78$) received the lowest mean score in this dimension; while this suggests that relatively less training was provided for this purpose, it is still classified as “sufficient.”

Table 4.*Descriptive Statistics of Received and Needed Training in Purposes of Testing*

LTA Practices	Received Training			Needed Training		
	Mean	SE	SD	Mean	SE	SD
Giving Grades	3.19	.068	.677	1.86	.107	1.07
Finding out What Needs to be Learned/ Taught	3.32	.058	.584	1.94	.107	1.07
Placing Students onto Programs	2.89	.080	.803	2.14	.112	1.11
Testing Competence in a Language	3.15	.061	.609	2.17	.113	1.12
Identifying What Has Been Learned	3.28	.062	.621	1.98	.108	1.08
Measuring General Ability to Learn L2	3.01	.080	.798	2.31	.113	1.13
Awarding Final Certificates	2.78	.089	.894	2.35	.107	1.06

Note. The Authors.

Results related to perceived training needs for testing purposes revealed that teachers usually stated requiring additional education at the basic level, with average scores varying from 1.86 to 2.35. More specifically, the highest mean scores were observed in more formal assessment purposes, including Awarding Final Certificates ($M = 2.35$), Measuring General L2 Learning Ability ($M = 2.31$), and Testing Competence in a Language ($M = 2.17$). However, teachers seem more comfortable with testing goals that are frequently adopted in regular educational environments, such as Giving Grades ($M = 1.86$), Finding out What Needs to be Learned/taught ($M = 1.94$), and Identifying What has been Learned ($M = 1.98$).

EFL Teachers' Received and Needed Training in Content and Concept of LTA

The Table 5 presents the findings of the descriptive statistics for the training received and required in Content and Concepts of LTA as perceived by in-service EFL teachers. With mean scores for the majority of items above 3.0, the overall training level for this dimension was sufficient. Testing Integrated Skills ($M = 3.26$), which is categorized as “Advanced” training, received the highest score. Testing Pronunciation ($M = 3.19$), Different Test Items for Vocabulary Knowledge ($M = 3.16$), and Testing Reading ($M = 3.13$) follow afterwards, all of which are located within the higher limit of the sufficient category. On the other hand, it was observed that teachers had received comparatively less training in some LTA concepts, despite still remaining at the “Sufficient” level.

The lowest mean scores were obtained in the use of statistics for assessment quality ($M = 2.91$), alternative assessment ($M = 2.92$), and practicality ($M = 2.93$), respectively. The findings also showed that the amount of education in language skills—reading ($M = 3.13$), writing ($M = 3.10$), speaking ($M = 3.08$), and listening ($M = 3.07$)—was quite comparable,

suggesting that teachers are offered reasonably equivalent instruction across these skills during teacher education. Moreover, while fundamental test quality characteristics such as Reliability ($M = 3.00$) and Validity ($M = 3.03$) were still within the “Sufficient” range, they fell below language skills and types of items.

Table 5.

Descriptive Statistics of Received and Needed Training in Content and Concept of LTA

LTA Practices	Received Training			Needed Training		
	Mean	SE	SD	Mean	SE	SD
Testing Reading	3.13	.065	.646	1.99	.117	1.16
Different Test Items for Reading	3.14	.064	.636	2.03	.115	1.15
Testing Listening	3.07	.070	.700	2.08	.113	1.12
Different Test Items for Listening	3.10	.073	.732	2.10	.111	1.11
Testing Speaking	3.08	.069	.692	2.18	.112	1.12
Different Test Items for Speaking	3.06	.071	.708	2.20	.110	1.10
Testing Writing	3.10	.069	.689	2.16	.117	1.17
Different Test Items for Writing	3.12	.070	.700	2.22	.118	1.17
Testing Grammar	2.97	.074	.745	2.26	.114	1.14
Different Test Items for Grammar	2.97	.076	.758	2.29	.113	1.13
Testing Vocabulary	3.13	.073	.734	1.89	.109	1.09
Different Test Items for Vocabulary	3.16	.072	.721	1.96	.111	1.10
Testing Integrated Skills	3.26	.071	.705	1.90	.111	1.11
Testing Pronunciation	3.19	.079	.787	1.93	.112	1.12
Different Test Items for Pronunciation	2.94	.080	.802	2.26	.116	1.16
Practicality	2.93	.079	.795	2.25	.111	1.11
Reliability	3.00	.078	.778	2.28	.113	1.12
Validity	3.03	.070	.703	2.29	.113	1.13
Authenticity	2.97	.078	.784	2.24	.108	1.08
Washback	2.98	.075	.752	2.31	.113	1.12
Using Statistics for Assessment Quality	2.91	.068	.683	2.37	.107	1.07
Alternatives in Assessment	2.92	.081	.813	2.34	.111	1.11

Note. The Authors.

When examining training demands in this domain, teachers generally reported a basic level of further training in most contents of LTA, while only a few areas attained rates around the “Intermediate” category ($M = 1.89$ – 2.37). In line with the received training results, higher mean scores and frequency distributions indicated the greatest training requirements in Using Statistics for Quality of Assessment ($M = 2.37$), Alternatives in Assessment ($M = 2.34$), and Washback ($M = 2.31$). Consistently, perceived training needs at the basic level were observed in the LTA contents, where teachers indicated receiving training at a higher level, such as Testing Vocabulary ($M = 1.89$), Integrated Skills ($M = 1.90$), and Pronunciation ($M = 1.93$).

Notably, many participants reported no perceived requirement for additional training in those areas, reflecting teachers' confidence in fundamental assessment practices. An examination of main language skills revealed that training needs of teachers reflect the "Basic" level ($M = 1.89$ – 2.29), with some areas scoring relatively lower mean scores than others, particularly Testing Vocabulary ($M = 1.89$) and Integrated Skills ($M = 1.90$). In this regard, a similar pattern emerged across all language skill groups, indicating that the items related to creating various items for specific skills showed higher need for training than those for testing these skills themselves. For instance, the mean score for Testing Reading ($M = 1.99$) was slightly lower than that of Different Test Items for Reading ($M = 2.03$).

DISCUSSION

The purpose of this study was to examine the LAL levels of in-service EFL teachers and their perceived needs for professional development. The first research question investigated how in-service English teachers perceive the scope and extent of the training they have received in language assessment. The findings showed that EFL teachers generally thought they had received sufficient training in language assessment across all dimensions, with the highest levels found in testing purposes and comparatively lower levels in testing and assessment knowledge.

Although previous studies have reported moderate or inadequate levels of LAL (Mede & Atay, 2017; Vogt & Tsagari, 2014; Xu & Brown, 2017), this discrepancy may be partially explained by differences in measurement scales and interpretation criteria. This finding may also reflect the impact of recent assessment policy in Turkey, which has increased teachers' involvement in assessment procedures based on four language skills and may have enhanced their perceptions of assessment competence.

At a more detailed level, the findings demonstrate a particular pattern across sub-domains, with teachers indicating higher levels of preparation in regular and skill-oriented assessment procedures, including using pre-made exams, providing feedback, and determining learning outcomes and needs. In contrast, more specialized areas, such as alternative assessment, computerized assessment, certification, and statistics, showed comparatively lower levels of training. This tendency is consistent with previous studies employing the same or similar instruments (Ballıdağ & İnan-Karagül, 2021; Kvasova & Kavytska, 2014; Mede & Atay, 2017; Sarıyıldız-Canlı, 2023; Vogt & Tsagari, 2014), which similarly found that teachers felt

more competent in routine classroom practices than in the technical and less frequently applied aspects of assessment. This pattern suggests an imbalance across LAL dimensions, where application-oriented skills develop more readily than conceptual and technical knowledge. A possible explanation is that pre-service assessment courses tend to emphasize routine classroom practices while devoting less attention to more specialized and technically demanding aspects of assessment.

The second research question aimed to identify teachers' perceived professional development needs across different dimensions of language assessment. The findings revealed that teachers generally expressed a basic-level need for further training across all dimensions despite perceiving their prior education as sufficient. This finding is largely consistent with previous studies reporting continuing professional development needs in language assessment (Ballıdağ & İnan-Karagül, 2021; Hasselgreen et al., 2004; Kvasova & Kavytska, 2014; Mede & Atay, 2017; Sarıyıldız-Canlı, 2023; Vogt & Tsagari, 2014). Among the language assessment domains, the highest need was observed in classroom-focused assessment. In particular, teachers reported relatively greater needs in areas such as proficiency, placement, aptitude, and diagnostic testing, as well as the English Language Portfolio. These findings are also consistent with previous research highlighting stronger demands for specialized language tests and portfolio assessment.

Additionally, relatively high levels of need were identified in computer-based assessment, statistics, and alternative assessment methods, indicating that many teachers may feel insufficiently prepared to meet current assessment expectations and emerging assessment practices. Conversely, lower levels of need were reported in routine classroom assessment practices and commonly assessed language skills, which may reflect teachers' greater familiarity and experience in these areas. Overall, the findings reveal that although in-service EFL teachers generally perceive their prior training in language assessment as sufficient, they continue to report a need for additional professional development across several areas of language assessment. Higher levels of need were particularly evident in more specialized domains, including alternative assessment, computer-based assessment, statistics, and specialized language tests. Taken together, these findings suggest that perceived levels of training do not necessarily correspond to preparedness for assessment practice, highlighting the importance of more targeted and practice-oriented professional development opportunities.

FINAL CONSIDERATIONS AND PEDAGOGICAL IMPLICATIONS

The findings of the current research highlight important implications for language assessment training in both pre-service and in-service education. The gap between teachers' perceived training levels and their continuing need for professional support suggests that current teacher education programs may not sufficiently promote in-depth and applicable assessment skills.

Accordingly, more systematic and practice-oriented training that enables teachers to engage in real-world assessment activities should be prioritized over the conceptual presentation of assessment content. In particular, greater emphasis should be placed on areas in which teachers reported higher training needs, including specialized and computerized test design, alternative assessment, and statistics.

Furthermore, assessment training should be viewed as an ongoing process and supported through continuous professional development opportunities aligned with changing educational and assessment policies. Such programs should also incorporate digital competencies and contemporary assessment approaches to better address the evolving needs of Turkish EFL teachers.

REFERENCES

- Arıcı, M. (2022). *Exploring the relationship between EFL teachers' language assessment literacy and their reflective teaching* [Unpublished master's thesis, Başkent University].
- Bağışlayıcı, A. (2023). *Designing an online language assessment training course for improving language assessment literacy of secondary school EFL in-service teachers in Mersin* [Unpublished doctoral dissertation, Yeditepe University].
- Bailey, K. M., & Brown, J. D. (1996). Language testing courses: What are they? In A. Cumming & R. Berwick (Eds.), *Validation in language testing* (pp. 236–256). Multilingual Matters.
- Ballıdağ, S., & İnan-Karagül, B. (2021). Exploring the language assessment literacy of Turkish in-service EFL teachers. *Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 24(45), 73–92. <https://doi.org/10.31795/baunsobed.909953>
- Bayram, M. (2026). *An investigation of language assessment literacy levels and training needs of Turkish EFL in-service teachers* [Unpublished doctoral dissertation, Hacettepe University].
- Borg, S. (2015). *Teacher cognition and language education: Research and practice*. Bloomsbury Academic.
- Brindley, G. (2001). Language assessment and professional development. In C. Elder, A. Brown, K. Hill, N. Iwashita, T. Lumley, T. McNamara, & K. O'Loughlin (Eds.), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 126–136). Cambridge University Press.
- Brookhart, S. M. (2001). *The standards and classroom assessment research* [Conference presentation]. Annual Meeting of the American Association of Colleges for Teacher Education, Dallas, TX, United States.
- Brown, J. D., & Bailey, K. M. (2008). Language testing courses: What are they in 2007? *Language Testing*, 25(3), 349–383. <https://doi.org/10.1177/0265532208090157>
- Bulut, N. (2024). *Different stakeholders' localized conceptualization of technology-enhanced language assessment literacy in higher educational context in Türkiye* [Unpublished master's thesis, Başkent University].
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11(1), 38–54. <https://doi.org/10.1080/09500789708666717>
- Clapham, C. (2000). Assessment and testing. *Annual Review of Applied Linguistics*, 20, 147–161. <https://doi.org/10.1017/S0267190500200093>
- Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing*, 28, 43–56. <https://doi.org/10.1016/j.asw.2016.03.001>

- Çağlayan, C. (2022). *An investigation into the relationship between language assessment literacy and self-efficacy beliefs of EFL instructors: An example of Ankara* [Unpublished master's thesis, Ufuk University].
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347. <https://doi.org/10.1177/0265532208090156>
- Dörnyei, Z., & Csizér, K. (2012). How to design and analyze surveys in second language acquisition research. In A. Mackey & S. M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 74–94). Wiley-Blackwell.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
- Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile: Issues in Teachers' Professional Development*, 20(1), 179–195. <https://doi.org/10.15446/profile.v20n1.62089>
- Han, T., & Kaya, H. İ. (2014). Turkish EFL teachers' assessment preferences and practices in the context of constructivist instruction. *Journal of Studies in Education*, 4(1), 77–93. <https://doi.org/10.5296/jse.v4i1.4873>
- Hasselgreen, A., Carlsen, C., & Helness, H. (2004). *European survey of language and assessment needs report: Part one – General findings*. EALTA.
- Hatipoğlu, 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 Journal*, 13, 40–51.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402. <https://doi.org/10.1177/0265532208090158>
- Inbar-Lourie, O. (2017). Language assessment literacy. In E. Shohamy, I. G. Or, & S. May (Eds.), *Language testing and assessment* (pp. 1–14). Springer.
- 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/10.1177/0265532213480334>
- Kaya, T. (2020). *Exploring language assessment literacy of EFL instructors in language preparatory programs* [Unpublished master's thesis, Bahçeşehir University].
- Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the Language Assessment Literacy Survey. *Language Assessment Quarterly*, 17(1), 100–120. <https://doi.org/10.1080/15434303.2019.1674855>
- Kvasova, O., & Kavytska, T. (2014). The assessment competence of university foreign language teachers: A Ukrainian perspective. *Language Learning in Higher Education*, 4(1), 159–177. <https://doi.org/10.1515/cercles-2014-0010>

- Levi, T., & Inbar-Lourie, O. (2020). Assessment literacy or language assessment literacy: Learning from the teachers. *Language Assessment Quarterly*, 17(2), 168–182. <https://doi.org/10.1080/15434303.2019.1692347>
- 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 11–15). *Measuring teachers' knowledge and application of classroom assessment concepts: Development of the Assessment Literacy Inventory* [Conference presentation]. Annual Meeting of the American Educational Research Association, Montreal, Canada.
- Ölmezer-Öztürk, E., & Aydın, B. (2019). Investigating language assessment knowledge of EFL teachers. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 34(3), 602–620. <https://doi.org/10.16986/HUJE.2018043465>
- Öz, S., & Atay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *ELT Research Journal*, 6(1), 25–44.
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381–402. <https://doi.org/10.1177/0265532213480337>
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, 48(1), 4–11. <https://doi.org/10.1080/00405840802577536>
- Rogers, W. T., Cheng, L., & Hu, H. (2007). ESL/EFL instructors' beliefs about assessment and evaluation. *Canadian and International Education*, 36(1), 39–61. <https://doi.org/10.5206/cie-eci.v36i1.9088>
- Sarıyıldız, G. (2018). *A study into language assessment literacy of preservice English as a foreign language teachers in Turkish context* [Unpublished master's thesis, Hacettepe University].
- Sarıyıldız-Canlı, G. (2023). *Language assessment literacy level and needs of in-service EFL teachers in Turkey* [Unpublished doctoral dissertation, 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/10.1177/0265532213480128>
- Shim, K. N. (2009). *An investigation into teachers' perceptions of classroom-based assessment of English as a foreign language in Korean primary education* [Unpublished doctoral dissertation, University of Exeter].
- Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Pre-service teachers' assessment literacy. *Journal of Science Teacher Education*, 22(4), 371–391. <https://doi.org/10.1007/s10972-011-9231-6>
- Spolsky, B. (1995). *Measured words: The development of objective language testing*. Oxford University Press.

- Stabler-Havener, M. L. (2018). Defining, conceptualizing, problematizing, and assessing language teacher assessment literacy. *Studies in Applied Linguistics & TESOL*, 18(1), 1–22. <https://doi.org/10.7916/salt.v18i1.1195>
- Stiggins, R. J. (1991). Relevant classroom assessment training for teachers. *Educational Measurement: Issues and Practice*, 10(1), 7–12. <https://doi.org/10.1111/j.1745-3992.1991.tb00171.x>
- Stiggins, R. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77, 238–249.
- Stiggins, R. J. (2008). *Assessment manifesto: A call for the development of balance assessment systems*. The ETS Assessment Training Institute.
- Stiggins, R. (2014). *Revolutionize assessment: Empower students, inspire learning*. Corwin Press.
- Sultana, N. (2019). Language assessment literacy: An uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, 9(1), 1. <https://doi.org/10.1186/s40468-019-0077-8>
- Şahin, S. (2019). *An analysis of English language testing and evaluation course in English language teacher education programs in Turkey: Developing language assessment literacy of pre-service EFL teachers* [Unpublished doctoral dissertation, Middle East Technical University].
- Şükür, M. (2023). *The description of pre-service EFL teachers' language assessment literacy development process* [Unpublished doctoral dissertation, Gazi University].
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403–412. <https://doi.org/10.1177/0265532213480338>
- Tian, W., Louw, S., & Khan, M. K. (2021). Covid-19 as a critical incident: Reflection on language assessment literacy and the need for radical changes. *System*, 103, 102682. <https://doi.org/10.1016/j.system.2021.102682>
- Tsagari, D., & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges, and future prospects. *Papers in Language Testing and Assessment*, 6(1), 41–64. <https://doi.org/10.58379/UHIX9883>
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of European study. *Language Assessment Quarterly*, 11(4), 374–392. <https://doi.org/10.1080/15434303.2014.960046>
- Xie, Q., & Tan, S. (2019). Preparing primary English teachers in Hong Kong: Focusing on language assessment literacy. *Journal of Asia TEFL*, 16(2), 653–673. <https://doi.org/10.18823/asiatefl.2019.16.2.14.653>
- Xu, Y., & Brown, G. T. (2017). University English teacher assessment literacy: A survey-test report from China. *Papers in Language Testing and Assessment*, 6(1), 133–158. <https://doi.org/10.58379/UZON5145>

- Yan, X., & Fan, J. (2021). “Am I qualified to be a language tester?” Understanding the development of language assessment literacy across three stakeholder groups. *Language Testing*, 38(2), 219–246. <https://doi.org/10.1177/0265532220929924>
- Yastıbaş, A. E., & Takkaç, M. (2018). Understanding language assessment literacy: Developing language assessments. *Journal of Language and Linguistic Studies*, 14(1), 178–193.
- Yılmaz, F. (2021). *An investigation of the psychometric properties of a language assessment literacy measure* [Unpublished master’s thesis, Middle East Technical University].
- Zhang, L., Zheng, Y., & Yang, D. (2026). Exploring Chinese university English teachers’ language assessment literacy: A mixed-method study. *Asia Pacific Journal of Education*, 46(1), 322–340. <https://doi.org/10.1080/02188791.2024.2354684>

CRediT Author Statement

- ☐ **Acknowledgments:** This research is based on the Ph.D. dissertation of the first author, titled *An Investigation of the Language Assessment Literacy Levels and Training Needs of Turkish EFL In-Service Teachers* (2026), which is supervised by the second author at the Graduate School of Educational Sciences, Hacettepe University.
 - ☐ **Funding:** The development of this research did not receive financial support from any funding agency.
 - ☐ **Conflicts of interest:** None.
 - ☐ **Ethical approval:** Ethical approval was granted by the Hacettepe University Institute of Educational Sciences Research Ethics Committee. Participation was voluntary, and informed consent was obtained from all participants.
 - ☐ **Data and material availability:** The data sources are publicly available on the respective websites.
 - ☐ **Authors' contributions:** All authors contributed equally to the article.
-

Processing and editing: Editora Ibero-Americana de Educação
Proofreading, formatting, normalization and translation

