Research Article

A comparison of Turkish and European English language teachers' language assessment knowledge levels and perceptions

Samet Fındıklı [ID][1*], **Kağan Büyükkarcı** [ID][2]

[1]Republic of Türkiye Ministry of National Education, Isparta, Türkiye
[2]Süleyman Demirel University, Faculty of Education, Department of Foreign Language Education, Isparta, Türkiye

Abstract: Language assessment knowledge, the capacity of language instructors to skillfully design, construct, and assess language evaluations, is pivotal for effective language education. This study investigates the language assessment knowledge, encompassing both general and skill-specific aspects, of in-service language educators from Europe and Türkiye. The primary objective is to contrast the language assessment knowledge of these two groups, highlighting potential differences in their assessment knowledge in terms of general and four language skills. Employing a mixed-methods approach, data were gathered sequentially via quantitative scale and qualitative online interviews. A total of 94 language teachers, 48 from Turkey and 46 from diverse European countries took part in this research. They completed the Language Assessment Knowledge Scale, and eight instructors engaged in semi-structured online interviews. The participants were selected using convenience sampling. The results indicated that while both groups scored above the average and were considered assessment literate, European language teachers had a significantly higher level of LAK compared to Turkish language teachers. This suggests that European teachers possess greater proficiency and competence in language assessment, potentially influencing the quality of the assessments they create and assess. Considering the importance of assessment knowledge mentioned in numerous studies, despite the limited sample size of this study, its results are important for the professional development of language educators. These outcomes can inform the development of teacher training programs, particularly for Turkish educators. The Ministry of National Education may consider prioritizing assessment-related subjects, such as assessing the four language skills, in future in-service teacher training initiatives.

1. INTRODUCTION

Assessment is accepted as the engine that drives learning (Cowan, 1998). Indeed, education and assessment are intertwined and indispensable units for each other. Although assessment and testing are mostly seen as scoring tools about how much learning has taken place in the classroom (Giraldo, 2018), they are also invaluable feedback that will guide the course of education (Mertler & Campbell, 2005). Especially in language education, which includes four

---

skills (reading, writing, speaking, and listening), the scope expands considerably, and the assessment knowledge level of the teachers is of great importance in terms of accurately gauging the education given and increasing the quality of the education shaped by the feedback from the assessment (Hughes, 2003; Malone, 2011; Popham, 2011; Stiggins, 1995).

While Popham (2011) agrees on the necessity of assessment literacy, he opposes what he perceives as the existing definitions in teachers' minds. According to Popham, assessment literacy is not solely "knowledge about educational tests and their roles," nor is it "the technical skills needed to construct and evaluate educational tests," or the ability "to calculate means, standard deviations, and correlation coefficients" (p. 267). Instead, he redefines assessment literacy as an individual's understanding of fundamental assessment concepts and procedures that are likely to influence educational decisions (Popham, 2011). This new definition serves as a reminder to educators that assessment not only measures but can also influence the course of education by providing feedback. Similarly, Boyles (2005) argues that "teachers and administrators need the necessary tools for analyzing and reflecting upon test data to make informed decisions about instructional practice and program design" (p. 18).

Having a better understanding of assessment procedures can have positive impacts on the quality of education, as stated by Malone (2011) who believes that "language assessment and language teaching go hand in hand. The best teaching involves high-quality assessment practices, and great assessment provides positive washback to the teaching and learning process" (p. 2). Thus, in order to apply successful assessment procedures in their classrooms and programs, educators need to have a strong foundation in assessment literacy, as emphasized by Malone (2011). Similarly, Giraldo (2018) argues that selecting, designing, and evaluating valid assessments is essential for achieving positive outcomes in learning and teaching. Furthermore, according to Büyükkarcı (2014), the systematic nature of assessment provides teachers with the opportunity to improve their teaching and provide the best learning experience for their students. This claim is supported by Cheng and Fox (2017), who noted that assessment plays a critical role in checking on learning and providing important information to teachers.

1.1. Assessment, Testing, and Evaluation

Assessment, testing, and evaluation in education play a crucial role in determining students' learning outcomes. While learning often leads to observable changes in performance, it is essential to recognize that learning is not always directly observable, as noted by Colby (2010). To bridge this gap, various methods and techniques are employed to measure unobservable behavioral changes, helping educators identify areas of mastery and improvement in learners (Douglas, 2009).

It is important to distinguish between the terms assessment, evaluation, and testing, as they are frequently used interchangeably. Assessment, defined by Coombe (2018), involves measuring an individual's performance to infer their abilities and provide feedback on their development. This process includes various methods such as tests, quizzes, and observations to gauge student learning (Brown, 2000; Rogiers, 2014). Assessment can be further categorized into formal and informal assessments. Informal assessments rely on observation and lack standardized rubrics, while formal assessments use standardized instruments and exams (Coombe, 2018). Both serve different purposes and have their advantages and disadvantages. There are also different assessment types, including diagnostic, self, peer, formative, and summative assessments, depending on their purpose and application.

Testing, as described by Nagai et al. (2020), is a specific type of assessment that involves formal tasks graded to gauge learners' language abilities. Tests are tools used to measure performance or knowledge, and they are designed with specific goals to draw desired conclusions about a student's abilities (Green, 2013; Bachman, 2004; Heaton, 1989).

Evaluation, a broader concept, involves the systematic gathering of information to make decisions (Bachman, 1990). It encompasses assessing program components, methods, or results to determine if they meet predetermined standards or objectives (Mohan, 2022). Evaluation also extends to assessing students, teachers, and curriculum effectiveness in relation to established goals.

In summary, assessment, testing, and evaluation serve distinct purposes in education, with assessment focusing on measuring individual performance and providing feedback, testing concentrating on formal tasks to gauge abilities, and evaluation encompassing a broader process of gathering information to make decisions about educational programs and outcomes.

1.2. Language Assessment Literacy

Language Assessment Literacy (LAL) is an important component of assessment literacy for language teachers, as it involves their conceptual knowledge and competence in testing, assessment, and evaluation. LAL is defined by Inbar-Lourie (2017) as "the essential knowledge, skills, and principles that stakeholders involved in assessment activities must master in order to perform assessment tasks effectively." Teachers devote a significant portion of their instructional time to assessment tasks, making it critical for them to be equipped with LAL skills and knowledge. DeLuca et al. (2015), Gotch and French (2014), and Siegel and Wissehr (2011) all highlight the significance of teacher preparation in assessment, covering topics such as test item creation, administration, evaluation, analysis, statistics, and reporting.

Davies (2008) emphasizes the importance of integrating skills, knowledge, and principles into teaching, whereas Scarino (2013) proposes integrating specialized knowledge of language assessment with an understanding of the interconnectedness of language, culture, and learning. According to Popham (2011), educators' assessment literacy influences their ability to make informed educational decisions, and Wiliam (2011) emphasizes the potential of integrating assessment with instruction to improve student engagement and learning outcomes.

Figure 1. *AL/LAL stakeholders (Taylor, 2013, p. 409).*
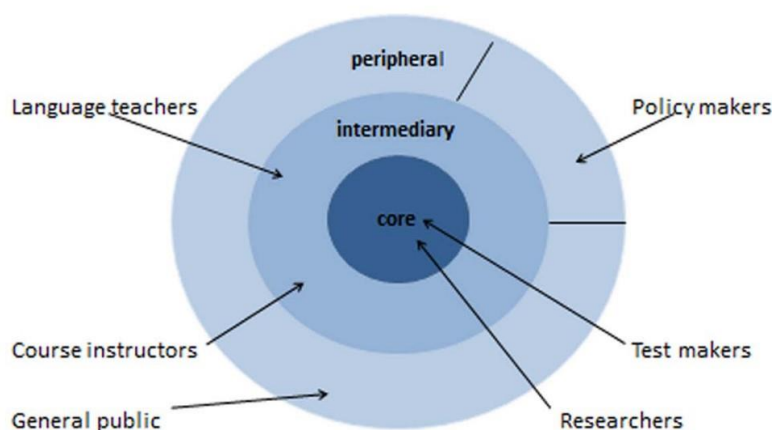
The literature on language assessment literacy addresses the question of which stakeholders should be literate in language assessment. Taylor (2013) proposes varying levels of assessment literacy based on the roles and responsibilities of stakeholders. As can be seen in Figure 1, researchers and test developers are regarded as the core group, necessitating a thorough understanding of assessment theory, technical expertise, and moral principles. Course instructors and language teachers are at the intermediate level, as they require practical expertise for test development while putting less emphasis on theory or ethical principles. Policymakers and the general public are in the outermost circle, where a basic understanding of test instrument characteristics and score significance suffices for decision-making.

1.3. Statement of the Problem

Assessment in language education encompasses testing and evaluation methods (Clapham, 2000). Language teachers are responsible for various assessment processes, including preparation, administration, evaluation of assessment tools, feedback provision, and informal observations (Ölmezer-Öztürk & Aydın, 2019). However, there is ongoing debate regarding whether language teachers receive adequate education and training to fulfill these responsibilities.

Assessment results help identify areas of weakness in language knowledge, determine students' needs, and evaluate the effectiveness of teaching (Harding & Kremmel, 2016). Good assessment practices also enhance teaching quality and student learning outcomes (Jannati, 2015). Moreover, assessment benefits students by identifying areas needing improvement, fostering self-assessment skills, and preparing them for high-stakes standardized tests (Thomas et al., 2004).

Language assessment literacy is essential for a teacher's professional development, and it requires both theoretical knowledge and practical implementation (Inbar-Lourie, 2008). A lack of adequate training in language assessment can lead to inadequate assessment practices and hinder student progress (Giraldo, 2021). Numerous researchers emphasize the importance of language assessment literacy for language teachers (DeLuca & Klinger, 2010; Fulcher, 2012; Harding & Kremmel, 2016; Lam, 2015; Malone, 2011; Scarino, 2013; Shepard, 2000; Siegel & Wissehr, 2011; Taylor, 2009). However, there is no consensus on specific competencies for language teachers in this area.

Fulcher (2012) notes that despite significant developments since the 1990s, language assessment literacy is still in its early stages. Recent research has highlighted the need for language teachers to receive adequate training in language assessment (Lam, 2015; Sarıyıldız, 2018; Sevimel-Şahin, 2019; Sevimel-Şahin & Subaşı, 2021; Tamerer, 2019; Wardani et al., 2021; Yetkin, 2015). Studies have focused on pre-service teachers, university-level English instructors, and in-service teachers, examining their assessment literacy levels, training needs, and perceptions.

Most existing studies on language assessment knowledge have been regional, focusing on specific geographic areas. Understanding regional differences in assessment practices is crucial, as cultural orientations and learner preferences can influence language assessment effectiveness (Krajka, 2019). Furthermore, the EF English Proficiency Index 2022 report highlights lower English proficiency levels in Türkiye compared to other European countries (EF EPI, 2022). This raises questions about potential links between language teachers' assessment knowledge and variations in proficiency levels. Further investigation into assessment practices among Turkish and European language teachers is needed to understand disparities and improve language education.

Language assessment is a significant part of education in Türkiye and Europe, with teacher education programs typically including coursework and practical training in assessment. These programs cover assessment principles, types, validity, reliability, and fairness. Findings from this study will shed light on the strengths and weaknesses of language teacher education programs, contributing to improved language education practices in Türkiye and Europe.

1.4. Purpose of the Study

By examining the language assessment knowledge of in-service language teachers, with a particular focus on Türkiye, this study aims to fill a significant gap in the literature. The main goal is to gauge these teachers' levels in language assessment and then compare it with that of their counterparts in European countries. By doing this, the study hopes to identify any potential variations in assessment procedures among language teachers from various countries and investigate how they may affect assessment knowledge. The study also compares and examines

the assessment abilities and knowledge of language teachers while taking into account cultural differences and standardization policies. This thorough investigation of language assessment practices will add to the body of knowledge already available on LAL by providing insightful information on the particular difficulties and variations that language teachers face in their assessment practices in various contexts. The research has a clear focus on both general and skill-based assessment knowledge, which will provide detailed information on the knowledge levels of in-service teachers. The study will also contribute to the development of effective language assessment practices in Türkiye and other countries. Overall, the research is expected to provide valuable insights into the AL of in-service language teachers and inform the development of effective language assessment policies and practices.

At the same time, this research will seek answers to the following research questions;

1. What is the Turkish and European EFL teachers' level of language assessment knowledge (LAK) in assessing students' language skills in English?
2. Is there a significant difference between the general and skill-based language assessment knowledge levels of Turkish and European language teachers?
3. How do country and demographic factors such as years of experience, educational background, school level, completion of a testing course, and attendance of testing and assessment training influence the overall LAK level and its skill-based components?
4. What are the perceptions of Turkish and European EFL teachers about their classroom-based language assessment practices?

2. METHOD

2.1. Research Design

This cross-national comparative educational study examined the assessment knowledge of English language teachers in Türkiye and European countries using mixed methods research. This approach combined qualitative and quantitative data to gain a comprehensive understanding of language teachers' assessment practices and knowledge (Cohen et al., 2018). Specifically, the study utilized an explanatory sequential design (Creswell, 2014), collecting quantitative data initially and then qualitative data to further explain the quantitative findings. This mixed methods approach enables a deeper exploration of language teachers' assessment knowledge and techniques in different educational contexts (Fox, 2016). It is claimed that mixed methods research design helps researchers address a wider range of concerns regarding the complex phenomena that are the focus of applied linguistic studies, and language assessment studies in particular, by moving beyond paradigmatic polarity (Fox, 2016). Additional sub-models are included in a mixed methods research.

Figure 2. *Explanatory design: Follow-up explanations model. Adapted from Creswell and Clark, 2007, p. 72.*
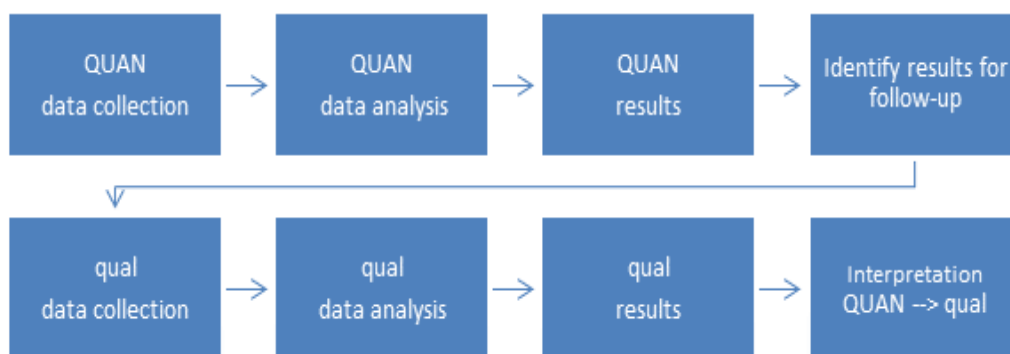

Figure 2 provides a visual representation of the quantitative and qualitative data gathered in the current study, which Creswell (2014) refers to as explanatory sequential design, aiming to gain

a deeper understanding of the assessment methods and expertise employed by language teachers.

2.2. Participants

This study collected data from two groups of participants: in-service English language teachers in Türkiye and those in European countries all of whom were actively working in middle and high schools. The primary aim was to compare the LAK of these language teachers. European countries were considered as a single group for analysis due to the complexity of handling each country individually (Lor, 2019).

The selection of these two groups enabled an investigation into potential differences in language assessment knowledge between them. Given that European countries generally exhibited higher levels of English language proficiency compared to Türkiye (EF EPI, 2022), it is hypothesized that European language teachers may possess higher assessment literacy, which could contribute to more accurate assessments and tailored teaching methods.

Table 1. *Countries of participants.*

Countries	N	Percent
Türkiye	48	51.1%
Italy	12	12.7%
Spain	12	12.7%
Romania	7	7.4%
Albania	4	4.3%
Bulgaria	4	4.3%
Germany	4	4.3%
Lithuania	3	3.2%
Total	94	100%

As can be seen in Table 1, the study maintained a balanced distribution of participants, with a total of 94 in-service language teachers, including 48 from Türkiye and 46 from European countries (Italy, Spain, Romania, Albania, Bulgaria, Germany, and Lithuania). All participants' countries were selected as countries where English is a foreign language, not a first or second language, and have similar language teaching objectives.

Table 2. *Crosstabulation of gender * BA program graduated from.*

		BA program Graduated From		
		ELT	Non-ELT	Total
Gender	Female	49	10	59
	Male	27	8	35
Total		76	18	94

Table 2 shows that 59 participants were female, while 35 were male. Furthermore, the participants predominantly held degrees in English Language Teaching (ELT) programs, with 76 of them having graduated in this field. Regarding school levels, 46 participants worked at middle school level, while the remaining 48 worked at high school level. These distributions provided a comprehensive view of language assessment knowledge among participants in both Türkiye and European countries.

Quantitative data collection initially involved convenience sampling, and snowball sampling was utilized to reach more participants, especially in European countries. The goal was to include language teachers from diverse European countries rather than focusing on a single country. The selection for qualitative interviews was based on volunteers from the quantitative

phase, with 8 participants representing a mix of Turkish and European teachers with varied years of experience and educational backgrounds.

2.3. Data Collection

The study was conducted in the 2022-2023 educational year and employed a mixed-methods approach to collect data, combining quantitative and qualitative research tools. For the quantitative aspect, the Language Assessment Knowledge Scale (Ölmezer-Öztürk & Aydın, 2018) was utilized to assess the language assessment knowledge of English as a Foreign Language (EFL) teachers. The Language Assessment Knowledge Scale (LAKS) underwent a rigorous development process involving expert review, teacher feedback, and validation by ELT and assessment experts. Following this, a pilot test with 50 teachers revealed issues of response consistency and participant engagement, prompting further refinement. Five experts then carefully evaluated each item and retained only those deemed fundamental for language teachers' assessment knowledge. As a result, the scale was pared down to 60 items, distributed across reading, listening, writing, and speaking constructs, representing a refined and validated version ready for wider implementation among language teachers. It included two main sections: demographic information and assessment knowledge questions. Participants were presented with 60 questions related to assessing reading, listening, writing, and speaking skills, to which they responded with "true," "false," or "don't know." In comparison to the original development process of the Language Assessment Knowledge Scale (LAKS), wherein a Cronbach's alpha coefficient of $\alpha= .91$ was reported, the current study yielded a coefficient of $\alpha= .768$. This discrepancy in reliability estimates may stem from differences in sample characteristics, testing conditions, or other methodological factors. It is important to note that reliability estimates can vary across different study populations and contexts. While the coefficient obtained in this study remains within an acceptable range, caution should be exercised when interpreting the scale's reliability in the specific context of this investigation.

The scale was converted into an online version using Google Forms. A combination of convenience and snowball sampling techniques was employed by the researcher to reach participants in both Türkiye and European countries through personal networks and contacts in the field of English language teaching. Participants completed the online form, and no personal information was required. However, participants were given the option to volunteer for the qualitative part of the study by providing their email addresses for further contact.

In the qualitative phase, individual semi-structured interviews were conducted with willing language teachers. These interviews followed an open-ended format, allowing participants to respond freely to a set of nine questions adapted from Jannati's study (2015). Semi-structured interviews provide a framework for exploration while allowing participants to express their perspectives in their own terms (Cohen et al., 2018). This qualitative approach complemented the quantitative data, offering deeper insights into participants' viewpoints and attitudes toward language assessment. Interviews were conducted using Zoom, with participants' consent for recording. For Turkish participants, interviews were conducted in Turkish, transcribed, and then translated into English. European participants were interviewed in English, and the interviews were transcribed into text format.

2.4. Data Analysis

The data analysis phase encompassed both quantitative and qualitative methods. In quantitative analysis, data were transferred to SPSS 26.0 for analysis. LAK levels were determined by participants' correct answers to more than half of the questions (30 out of 60 questions). As the developers of the scale applied in their own research, participants who gave 30 or more correct answers were accepted as assessment literate. Participants who gave correct answers below 30 were accepted as inadequate in terms of assessment knowledge. Inferential statistics were used to compare participants' LAK levels based on various factors such as country, gender, educational level, and years of experience.

In qualitative analysis, interview data were processed using MAXQDA Analytics Pro 2020 software. Hypothesis-related code schemes were created, and interview responses were numbered for the organization. Data were selectively included to support quantitative results and provide additional context. Anonymity was maintained by assigning code names to participants. Rigorous research techniques were employed to ensure validity and reliability, including having the interview questions analyzed by experts in the field, choosing interviewees from diverse countries, recording the interviews, and using open-ended questions to encourage participants to answer freely rather than just yes or no answers, member checking, and continually comparing data with the codes.

3. FINDINGS

3.1. Quantitative Findings

The quantitative phase of the study focused on assessing the general and skill-based Language Assessment Knowledge levels of English language teachers. Initially, an analysis was conducted to assess the LAK levels of participating teachers. Subsequently, a comparison was made between the general and skill-based LAK levels of teachers in Türkiye and Europe, organizing them into two distinct groups. Additionally, the study examined whether the demographic factors included in the research scale had a significant impact on the LAK levels of EFL teachers.

To determine the suitability of statistical tests for further analysis and comparisons between Türkiye and Europe, tests assessing the normality assumptions of the LAK level variable were performed.

Table 3. *Results of the tests of normality for general LAK level.*

		Kolmogorov-Smirnov			Shapiro-Wilk		
	Country	Statistic	*df*	Sig.	Statistic	*df*	Sig.
LAK level	Türkiye	.088	48	.200	.988	48	.911
	Europe	.123	46	.079	.957	46	.087

Table 3 displays the outcomes of these tests, including the Kolmogorov-Smirnov and Shapiro-Wilk tests, for both country groups. The *p*-values from these tests, exceeding .05 for both Türkiye and Europe, indicated that there was no compelling evidence to suggest substantial deviations from normality in the LAK level variable. Consequently, parametric tests assuming normality, such as the t-test or ANOVA, were employed to compare LAK levels between the two country groups.

3.1.1. *General and skill-based LAK levels of the participants*

Table 4 presents the general LAK levels of all EFL teachers who participated in the study. Descriptive statistics, including the mean, standard deviation, minimum, and maximum values, were used to determine the participants' LAK levels.

Table 4. *General LAK level of EFL teachers.*

N	Minimum	Maximum	Mean	Std. Deviation
94	7	51	32.11	7.33

As shown in Table 4, the 94 English language teachers who participated in the study answered an average of *X*= 32.11 questions correctly out of the 60-question scale. The lowest number of correct answers was 7, and the highest number of correct answers was 51. Additionally, the skill-based LAK levels of the participants were also analyzed, in addition to their LAK levels.

Table 5. *Skill-based LAK levels of the participants.*

	N	Minimum	Maximum	Mean	Std. Deviation
Assessing Reading	94	3	14	9.24	2.04
Assessing Listening	94	0	15	7.20	2.62
Assessing Writing	94	0	13	7.41	2.47
Assessing Speaking	94	0	13	8.25	2.44

The research scale comprised 15 questions for each skill, allowing participants to score between 0 and 15 for each skill. According to Table 5, out of the 94 English language teachers who participated in the study, the highest mean score in skill-based analysis was obtained in reading assessment (X= 9.24). Among the four skills that constitute the English language, the skill with the lowest mean value among the questions asked was listening (X= 7.20).

A one-sample t-test was used to determine whether the score was significantly high. The lowest possible score on the scale is 0, and the highest possible score is 60. Therefore, 30 was selected as the reference point, which represents half of the total score.

Table 6. *One-sample t-test results of participants' general LAK level scores.*

Mean Diff.	df	t	p
2.11	93	2.799	.003*

* $p < .05$

Table 6 indicates that the mean difference (2.11) between all of the participants' mean scores on the scale (X= 32.11) and half of the maximum score (30) was statistically significant, which suggests that their overall LAK level is high.

After discovering that the mean scores of the participant teachers were significantly high, the same one-sample t-test was performed for each skill individually. However, this time, since there were 15 questions for each skill, 7.5 was used as the reference value. This approach aimed to determine whether there was a significant difference between the reference value of each skill and the score that teachers received for that skill.

Table 7. *Skill-based one-sample t-test results.*

	Mean diff.	Mean	df	t	p
Assessing Reading	1.74	9.24	93	8.28	< .001*
Assessing Listening	-0.30	7.20	93	-1.10	.274
Assessing Writing	-0.09	7.41	93	-0.33	.739
Assessing Speaking	0.75	8.25	93	2.99	.004*

*$p < .05$

Table 7 demonstrates that the mean difference (1.74) between the reference value (7.5), which was accepted as half of the total 15 points, and the mean score of the reading assessment skill (X= 9.24) indicated that the teachers' knowledge of measuring this skill was significantly high (p= < .001). A similar result was found for another skill, speaking, where the mean score for assessing speaking skills among the 94 teachers (X= 8.25) was slightly higher (0.75) than the reference score (7.5). The significance value (p= .004) suggests that the mean score of the participant teachers is also significantly high in evaluating this skill. However, the mean scores for assessing listening (X= 7.20) and assessing writing (X= 7.41) obtained by the teachers for the other two skills were slightly below the reference score. Based on the obtained data, it was found that the knowledge of the 94 participating teachers in the areas of assessing listening and writing was not significantly lower than the half scores, as indicated by the non-significant significance values (p= .274 and p= .739 respectively).

3.1.2. *A Comparison of the Turkish and European EFL teachers in terms of general and skill-based LAK*

The general LAK levels of two groups, Türkiye and Europe, were compared using an independent samples t-test, as shown in Table 8.

Table 8. *T-test results of General LAK levels by country of participation.*

Country	*N*	Mean	Std. Deviation	t-test *t*	*df*	*p*
Türkiye	48	30.64	4.88	-2.02	92	.046*
Europe	46	33.65	9.02			

*p< .05

It revealed a significant difference in general LAK levels between the two groups (*t*[92]=-2.02; *p*< .05). European participants (*X*=33.65, *SD*=9.02) demonstrated higher general LAK levels compared to Turkish participants (*X*=30.64, *SD*=4.88).

Additionally, the study assessed skill-based LAK levels in reading, listening, writing, and speaking. Each skill had a maximum score of 15, with 7.5 as the reference point for competence.

Table 9. *T-test results of Skill based LAK levels by country of participation.*

	Country	*N*	Mean	*SD*	t-test *t*	*df*	*p*
Assessing Reading	Türkiye	48	9.50	1.71	1.23	92	.220
	Europe	46	8.97	2.32			
Assessing Listening	Türkiye	48	6.47	1.92	-2.82	92	.006*
	Europe	46	7.95	3.04			
Assessing Writing	Türkiye	48	6.83	2.02	-2.37	92	.020*
	Europe	46	8.02	2.75			
Assessing Speaking	Türkiye	48	7.83	2.15	-1.72	92	.089
	Europe	46	8.69	2.66			

*p< .05

Table 9 presents the skill-based LAK levels for both groups. Participants from both groups demonstrated proficiency in reading and speaking skills, with no significant differences. In terms of assessing listening skills, European teachers displayed a mean score of *X*=7.95, while their Turkish counterparts exhibited a mean score of *X*=6.47 (*p*=.006). Additionally, concerning assessing writing skills, European language teachers attained a mean score of *X*=8.02, surpassing the mean score of *X*=6.83 achieved by Turkish teachers (*p*=.020).

In summary, European teachers generally exhibited higher LAK levels, especially in listening and writing skills, while Turkish teachers had a slight advantage in reading assessment. However, the differences in reading and speaking skills were not statistically significant.

3.1.3. *BA program graduated*

A comparative analysis of LAK levels between participants from Türkiye and Europe, based on their graduation from ELT or non-ELT BA programs, was conducted. Table 10 summarizes the findings.

Table 10. *T-test results according to BA program graduated.*

BA program graduated	Country	N	Mean	Std. Deviation	Sig.
	Türkiye	45	30.64	4.97	
ELT	Europe	31	34.14	8.11	.014[*]
	Total	76	32.18	6.65	
	Türkiye	3	30.66	4.04	
Non-ELT	Europe	15	32.06	10.80	.831
	Total	18	31.83	9.91	
Total		94	32.11	7.33	.856

*$p< .05$

For ELT graduates, Türkiye had a mean LAK level of X=30.64, Europe X=34.14, and the total X=32.18. The significant difference between Türkiye and Europe (p= .014) indicated variation.

Non-ELT graduates in Türkiye had a mean LAK level of X=30.66, Europe X=32.06, and the total X=31.83. No significant difference was found between Türkiye and Europe (p= .831). Overall, when considering both countries, the analysis revealed no significant difference in LAK levels between ELT and non-ELT graduates (p= .856).

3.1.4. *Testing course at university*

This analysis compared participants from Türkiye and Europe based on whether they took a testing course during their undergraduate studies, aiming to understand its impact on their LAK levels.

Table 11. *T-test results of LAK levels according to testing course at undergraduate education.*

Testing course at undergraduate education	Country	N	Mean	Std. Deviation	Sig.
	Türkiye	27	30.96	4.62	
Yes	Europe	29	35.27	8.36	.022*
	Total	56	33.19	7.10	
	Türkiye	21	30.23	5.30	
No	Europe	17	30.88	9.67	.796
	Total	38	30.52	7.46	
Total		94	32.11	7.33	.083

*$p< .05$

According to Table 11 for those who took the course, Turkish participants had a mean LAK level of X=30.96, European participants X=35.27, and the total X=33.19. A significant difference between the participants of Türkiye and Europe (p= .022) suggests the influence of the course.

Among those who did not take the course, Türkiye had X=30.23, Europe X=30.88, and the total X=30.52. No significant difference (p= .796) was observed between Türkiye and Europe in this group.

3.2. Qualitative Findings

The qualitative phase aimed to understand language assessment perceptions and practices among 4 Turkish and 4 European teachers. Interview questions adapted from Jannati (2005) explored their viewpoints and methods. The findings, divided into two sections, compare Turkish and European teachers' perspectives on language assessment (questions 1, 2, 4, 5, 8, 9) and their assessment methods (questions 3, 6, 7). This comparison revealed similarities and differences in their approaches to teaching and assessing students in different contexts.

3.2.1. *Findings about the EFL teachers' perceptions about language assessment*

An analysis of the perspectives shared by both Turkish and European participants revealed a range of opinions within the group. Regarding the need for assessment, several participants underscored its significance in monitoring student advancement, pinpointing areas requiring additional support, and providing constructive feedback to enhance teaching. For example, certain participants expressed:

> *"I believe we do need, and I believe we need assessment to inform our planning. As teachers, we need to know where our students are. That's what assessment for." (EU-4)*
>
> *"We need to prove that we are teaching English to children at some point. For this reason, we need to know whether children have learnt it or not. For this reason, there is a need, but under normal conditions, I think how much they can master the language is not fully measurable because it is very subjective." (TR-4)*

The fourth interview question, "Do you think students' scores represent what they have learned?", delves into the participants' perspectives on the relationship between students' scores and their actual learning outcomes. In response to this question, although they gave different answers and reasons, there was a consensus among the Turkish participants that grades do not represent what students have learned. Some responses of Turkish participants were:

> *"After grading the exams, I look at them and I say that some students got higher than the score they should have gotten. Both the difficulty level of the questions we ask students and our education system unfortunately do not measure children in a multidimensional way." (TR-1)*
>
> *"It definitely does not represent. We cannot say by looking at an exam grade on a paper that this student got a hundred means that he knows everything." (TR-2)*

However, the answers of the European participants varied. Some said that the grades represent students' language knowledge, while others disagreed. Some of the views of the European participants are as follows:

> *"If it's a reliable test and a valid test if it's well prepared, yes, the scores should represent what students have learned. Also, bearing in mind it's a flexible test, so it can be adjusted to the special needs of students." (EU-4)*
>
> *"For my local students or other international students, usually the scores represent exactly what they have learned. I don't know what the mystery about the Turkish students is. I really don't understand what's happening. You know, either they are shy, they don't know how to interact. My Turkish students Rümeysa and Betül don't speak any foreign language, but in writing they are excellent." (EU-3)*

Responses to the question, "How do you increase your knowledge about assessment?" provided insights into the participants' strategies for enhancing their understanding and competence in assessment practices. Turkish and European participants presented varying perspectives on their approaches to professional development and lifelong learning in this context.

> *"We invite many foreign speakers. They're mainly from universities, Oxford University or Cambridge University, it depends. They explain to us how to use the textbooks and how to give assessments. Formal or informal. How to provide uh, well, some exercises. Also, we have some video tasks... And I think we in schools have a group of teachers of foreign language teachers where we decide what to assess, how to assess and how many points we are going to give, so there must be an agreement among language teachers." (EU-1)*
>
> *"I tend to read around the topic and I also, as I said at the beginning, have attended a number of seminars, but these were more like 2-3 day conferences or trainings where there were interesting speakers. So, teachers were given the chance to voice their concerns, to discuss the problems they have in classrooms, and I found it really beneficial." (EU-4)*
>
> *"I didn't attend any workshops or anything like that, I tried to read a little bit about the subject at the time, but it became so branched and knotted somewhere, I can honestly say that I gave up... We discuss this with my friends all the time, let's say when the time comes, not all the time, but when the time comes, we talk about it. A small exchange of ideas, after a while, I mean, apart from that, there is nothing else, to be honest." (TR-3)*

3.2.2. *Findings about the EFL teachers' classroom-based assessment practices*

The participants were questioned about their approach to informing students about rubrics, focusing on the transparency and communication of assessment criteria. The data analysis revealed unanimous agreement among both Turkish and European participants that students should be informed about the assessment rubric. This shared perspective underscores the importance of transparency and providing students with clear guidance on the criteria used for evaluating their work. It reflects a common commitment to promoting fairness and enabling students to understand and meet the expected assessment standards. Sample responses from participants include:

> *"It's a must, and I always tell my students, never, never sit for an exam if you don't know, for example, what that exam includes in the sense what type of rubrics does it have? Because you know all kinds of exams, for example, maybe they include, or they want to test different things."* (EU-2)
>
> *"I think it would be helpful for children or students, adults, whoever they are, to be aware of those rubrics, to know what the goal of the person receiving instruction is and to draw their path accordingly."* (TR-3)

When language teachers were queried about the specific language skills or components they assess, a notable contrast emerged. All European language teachers indicated that they assess a foreign language as a comprehensive whole. In contrast, Turkish participants predominantly mentioned that they focus on teaching and assessing reading, grammar, and vocabulary. This divergence was attributed to the examination system in Türkiye. Here are some responses from European English language teachers:

> *"We cover all skills reading, writing, listening, and speaking. The main focus is on improving communication skills, so mostly oral skills are, I don't know, practiced during the semester, but the final examination is usually written exam. So, the writing component is also very important for them."* (EU-3)
>
> *"At first, I was focusing on speaking and listening, but since I realized that I was stealing their time, I don't evaluate them, I don't measure them, and I don't spend much time on them. I work more for the exam. Since these are not in the exam, I focus more on vocabulary as a language component."* (TR-2)
>
> *"They make a whole. Yes, this is what I said before. They are just like the fingers, for example, of one hand and in a way, if one of them does not work, the whole hand does not work properly."* (EU-2)

4. DISCUSSION and CONCLUSION

This study aimed to assess the language assessment knowledge of in-service language teachers in Türkiye and compare them with European counterparts. It utilized a mixed-methods approach, combining quantitative data from a knowledge scale and qualitative data from interviews. The findings were discussed in relation to research questions, implications for teacher education, and connections with existing research. Gaps in the literature were identified, suggesting areas for future research, and expanding the current knowledge in the field.

4.1. Discussion of the First Research Question

The purpose of the first research question was to find out the general and skill-based language assessment knowledge level of EFL teachers working in secondary and high schools in Türkiye and Europe. The results indicated that these teachers generally possessed a relatively high level of general knowledge about language assessment. However, when their skill-based assessment proficiency was analyzed, it was clear that although they performed exceptionally well when it came to assessing speaking and reading, they did not meet the reference score when it came to assessing writing and listening. This suggests that while teachers may have theoretical knowledge about various types of language assessment and their purposes, they may lack the practical skills required to construct valid and reliable assessments, particularly in listening and

writing areas. This underscores the importance of targeted professional development programs to enhance teachers' skill-based knowledge in language assessment, including the design, implementation, and evaluation of effective assessments.

The disparities in skill-based knowledge among language teachers can be attributed to the prioritization of reading and speaking skills in language teaching and assessment practices. These skills are often emphasized in curriculum, textbooks, and standardized tests, leading teachers to become more familiar and proficient in evaluating them. Conversely, listening and writing skills tend to receive less attention in educational settings, resulting in teachers having relatively less knowledge and experience in assessing these areas. Although EFL instructors acknowledge the significance of assessing oral skills, the findings of the study of Kim (2014) reveal a discrepancy between belief and practice. Despite their recognition of the importance of oral assessment, exams lack a dedicated speaking section, indicating that oral skills are not given as much importance in assessment practices. It is also worth noting that the findings of the current study align with a previous research conducted by Kırkgöz et al. (2018), which emphasized the underappreciation of listening and writing assessment among language teachers. The prevailing focus on reading and speaking skills is influenced by curriculum priorities and educational program objectives, with an emphasis on improving students' reading and speaking abilities (Altan, 2017).

Furthermore, the results of the study suggest that practical experience and on-the-job learning contribute significantly to teachers' assessment knowledge, reinforcing the idea that assessment literacy is primarily acquired through classroom practice rather than theoretical knowledge obtained during undergraduate education (Mertler, 2003). This distinction is evident when comparing the knowledge levels of in-service teachers in this study with those of pre-service teachers in another research (Çetin-Argün, 2020). The in-service teachers exhibited higher general LAL levels, likely due to their years of classroom experience. However, the study underscores the importance of enhancing teachers' understanding of assessment in writing and listening skills, as these areas have historically received less attention in teacher education programs. The findings of this study are consistent with various studies using the same assessment instrument, contributing to the reliability and validity of the results. Nevertheless, some contradictory findings in other studies highlight the complexity of the subject and the need for further research (Lam, 2015; Ölmezer-Öztürk & Aydın, 2018; Vogt & Tsagari, 2014; Xu & Brown, 2017).

4.2. Discussion of the Second Research Question

In this section, the second research question aimed to explore potential disparities in language assessment knowledge levels between Turkish and European language teachers, focusing on their general language assessment knowledge and skill-based language assessment knowledge. In aligning with the global trends in language education, the Turkish Ministry of Education has recently introduced a foreign language teaching program that bears striking similarities to those found in European countries (Turkish Ministry of Education, 2018). Emphasizing a communicative approach to assessment, the program underscores the importance of designing tasks that prioritize the practical application and production of language skills. This echoes the principles set forth by the Common European Framework of Reference for Languages (CEFR), a widely recognized framework that serves as a benchmark for language learning and assessment across Europe and beyond. By embracing these pedagogical principles, Turkey's language education program not only aligns with international standards but also fosters a learning environment that promotes effective communication and linguistic proficiency, mirroring the goals and objectives seen in European language education systems.

Analysis of the mean values revealed that European participants had higher scores than Turkish participants, with a statistically significant difference noted. When assessing skill-based knowledge, European respondents demonstrated higher mean values in all skills except for

assessing reading. This indicates that, on average, European language teachers possess greater LAK compared to their Turkish counterparts, highlighting a potential knowledge gap between the two groups. However, it is essential to note that these results are based on aggregate mean values and do not necessarily represent individual performances.

Several factors may contribute to the observed disparities in LAK levels between European and Turkish language teachers. Past research suggests that differences in educational systems and resources dedicated to language assessment practices across regions can play a role (Bonnet, 2007; Cheng et al., 2004; Jones & Saville, 2009; Vogt & Tsagari, 2014; Vogt et al., 2020). European countries, known for their well-established assessment frameworks such as CEFR and ample resources, may have an advantage, leading to the higher mean values among European respondents. In contrast, studies have highlighted challenges faced by Turkey, including limited availability of standardized assessments and inadequate resources allocated to language assessment (Büyükkarcı, 2016; Krajka, 2019; Mede & Atay, 2017), which may affect the performance of Turkish participants in these assessments. A research carried out by Ünlücan-Tosun and Glover (2020) found that Turkish language instructors expressed a lack of confidence in integrating CEFR levels into their classroom assessments. Additionally, they noted that course materials lacked sufficient guidance for effectively implementing the CEFR.

Additionally, the findings align with the study conducted by Çakır (2020), which examined language teachers' beliefs about assessment types, content, and skills across different countries. Çakır's study revealed that while beliefs about assessment types and content did not significantly differ, variations were observed in the reasons for utilizing classroom assessment among different countries. These findings, in conjunction with the present study, emphasize the influence of country-specific factors on language assessment practices and outcomes. Similarly, Cheng et al. (2004) conducted research across different countries and identified diverse assessment methods and procedures employed in ESL/EFL teaching and learning. This diversity in assessment approaches underscores the role of context and culture in shaping assessment practices in ESL/EFL education.

Although not statistically significant, Turkish participants showed superior performance in assessing reading skills. This could be attributed to the examination-oriented approach of the Turkish education system, which places a strong emphasis on reading comprehension skills, frequently assessed in high-stakes examinations (Hatipoğlu, 2010).

4.3. Discussion of the Third Research Question

The third research question sought to investigate the impact of the country and two demographic factors, educational background, and testing course at undergraduate education, on both the overall language assessment knowledge level and its skill-based components. The purpose of the investigation was to determine how these factors affect teachers' proficiency in language assessment. We can gain insights into the relationship between demographic features and LAK levels by examining differences across countries. The findings shed light on the extent to which country and specific demographic factors influence variations in language assessment knowledge, both at the general and skill-specific levels.

The study examined assessment knowledge and its relationship with the educational background of BA program graduates. Results show that participants who completed English Language Teaching (ELT) programs in Europe had significantly higher assessment knowledge levels compared to those in Turkey. However, there was no significant difference in assessment knowledge levels between participants from Turkey and Europe who completed non-ELT BA programs. This suggests that the country factor played a role in ELT program graduates' assessment knowledge but not in non-ELT program graduates. Similar studies by Genç et al. (2020) and Kaya and Mede (2021) found no significant difference in assessment literacy scores between ELT and non-ELT program language teachers. This suggests that individual factors like motivation, effort, and language proficiency may have a stronger influence on assessment

performance than the specific program type. Additionally, the study questions the effectiveness of assessment courses within ELT programs, aligning with Hatipoğlu's (2015) findings that despite extensive exposure to English language exams, ELT students had limited knowledge about testing in general and English language testing specifically.

The study found that the presence of a testing course during undergraduate education did not lead to a significant difference in assessment knowledge scores between language teachers who had taken the course and those who had not when considering the country factor. This suggests that having a testing course alone may not substantially impact language teachers' assessment knowledge levels, and several factors like course effectiveness, practical application, available support, and individual differences among teachers may be at play.

Furthermore, when focusing on participants who took the testing course, those from Europe had higher mean assessment knowledge levels compared to their Turkish counterparts. This discrepancy raises the possibility that the efficacy or nature of testing programs may differ between these areas, which could affect the assessment literacy of language instructors. According to Şahin (2009), a single LTA course is insufficient for adequately enhancing the assessment knowledge of prospective language educators to handle the demanding and crucial responsibility of consistently evaluating their students for both summative and formative assessment objectives.

These findings align with the need to reevaluate the role and effectiveness of testing courses in teacher training programs worldwide. A study conducted by Ölmezer-Öztürk and Aydın (2018) also found that having a separate testing course during BA degree education does not significantly impact the LAK levels of language teachers. Additionally, insights from Stiggins (1995) suggest that taking an educational testing and measurement course may not effectively prepare teachers for the practical realities of classroom life. This collective perspective underscores the importance of reevaluating the role of testing courses in equipping teachers with the necessary skills for assessment practices in real classroom contexts.

4.4. Discussion of the Fourth Research Question

When comparing Turkish and European EFL teachers' opinions on the value of assessment in language classes, it became clear that different contexts—cultural and educational—had an impact on their viewpoints. Assessment was emphasized by Turkish teachers as a means of understanding student learning and fulfilling exam-related requirements. Lam (2015) emphasizes how putting too much emphasis on exams can make learning less important. European educators understand the value of assessment in determining student proficiency and getting them ready for national exams. The impact of participants' prior experiences as language learners and teachers on their perceptions of assessment is noted by O'Loughlin (2006).

Turkish and European EFL teachers held opposing views on whether scores accurately represent students' learning. Turkish teachers expressed skepticism, citing environmental factors and exam-related stress as limitations in score accuracy. They pointed to cases of competent students struggling with exams. In contrast, most European teachers believed scores were accurate, highlighting potential regional differences in grading policies. This raises concerns about assessment practices aligning with the broader goals of communicative teaching approaches in language education (DeLuca et al., 2017; Gkogkou & Kofou, 2021).

The perspectives of Turkish and European EFL teachers revealed disparities in the skills assessed in their classes. Turkish educators prioritized reading and grammar assessments, considering them simpler and exam-relevant. In contrast, European teachers emphasized evaluating all language skills, particularly focusing on oral communication, aligning with the assessment for learning culture prevalent in Western countries (Xu & Brown, 2016). The variations can be attributed to the contrasting assessment cultures, with Western countries

emphasizing learning-oriented assessment and East Asian educational systems prioritizing high-stakes tests and rote memorization.

The study examined how Turkish and European EFL teachers approach enhancing their assessment knowledge. Turkish teachers emphasized the significance of in-service training for their professional development, and some recognized the value of academic courses, like master's degree programs, for deepening their understanding of assessment. In contrast, European teachers outlined diverse methods for expanding their assessment knowledge, including reading assessment books, attending workshops, seminars, and engaging in international conferences and partnerships. Herrera and Macias (2015) underscore the importance of integrating assessment literacy into language teacher education programs, emphasizing the need for continuous development and commitment to assessment knowledge among both novice and experienced educators.

4.5. Conclusion

In conclusion, the goal of this study was to reveal language teachers' assessment knowledge levels and compare Turkish in-service language teachers with teachers from other European countries in terms of their language assessment knowledge levels, and their in-class assessment practices. The results showed that, despite having a generally high level of knowledge about language assessment, EFL teachers in both country groups had varying skill-based levels. Teachers performed particularly well on the reading and speaking assessments, but less well on the listening and writing assessments.

Moreover, the results of this study not only confirmed the existence of varying levels of language assessment knowledge between Turkish and European language teachers but also revealed a notable disparity favoring the European group. With the exception of reading assessment, European participants showed higher mean values in both general and skill-based LAK. However, despite the observed differences in language assessment knowledge, it is noteworthy that no significant difference was found between the Turkish and European language teachers in terms of their assessing reading and speaking abilities. This finding suggests that, in these particular language skills, both groups demonstrated comparable levels of proficiency. On the other hand, when it came to assessing listening and writing skills, significant differences emerged, with the European teachers exhibiting higher expertise.

This study also sought to explore and compare the general and skill-based assessment knowledge levels of participants across different countries, taking into account various demographic factors. The findings revealed noteworthy differences in the areas of the participants' fields of study during their BA programs and, whether they had taken a testing course at university.

Based on the qualitative findings from the interviews, a clear difference in approaches to assessing the four language skills emerged between Turkish and European language teachers. Turkish teachers expressed concerns about national exams, causing them to focus solely on exam-oriented skills while ignoring the comprehensive assessment of all four skills. European teachers, on the other hand, prioritized communication skill development and recognized the importance of assessing all four language skills. This disparity can be attributed to different priorities for lifelong learning. When we examine the answers of the participants, it is clear that Turkish teachers placed little emphasis on acquiring new knowledge about their profession, whereas Europeans valued lifelong learning as a means of expanding their knowledge. Furthermore, Turkish teachers were skeptical regarding the representation of students' knowledge through grades, whereas Europeans saw grades as more important indicators of students' understanding. Despite these differences, Turkish and European teachers agreed on the importance of assessment, the use of rubrics, and various in-class assessment methods.

4.6. Limitations of the Study

The study has several limitations. Firstly, the online data collection method aimed to collect responses from a large number of English language teachers; however, only a small sample of language teachers from Türkiye and only seven countries in Europe completed the scale, which might not accurately represent the entire population of English language teachers in the study area. Secondly, the voluntary participation in the qualitative part of the research led to a limited number of interviews, potentially reducing the applicability and inclusiveness of the findings to a broader population. Furthermore, the online format of the scales could have resulted in a non-representative sample, and participants' English proficiency and interpersonal communication skills may have influenced the accuracy of their responses during the interviews.

Additionally, it is crucial to acknowledge that the education systems in diverse countries may vary significantly. These differences could impact the experiences and perspectives of English language teachers, introducing an additional layer of complexity to the interpretation of our findings. Despite our efforts to address these variations, it is important to interpret the study's outcomes with caution, given the low effect size observed, and to consider them within the context of the specific educational landscape in the study area.

4.7. Suggestions for Further Studies

The study's findings offer guidance for future language assessment research. Subsequent studies may explore specific factors, including cultural norms, educational frameworks, and institutional contexts, influencing the observed differences in language assessment knowledge between Turkish and European teachers. Expanding the sample size in future studies is crucial for broader applicability, allowing for a more comprehensive analysis of language assessment knowledge among English language teachers in both Turkish and European settings. Additionally, investigating students' perceptions and attitudes towards language assessment can inform tailored assessment design. Longitudinal studies could provide insights into the enduring impact of language assessment practices on students' language development and real-world language skill application, offering avenues for further research.

4.8. Pedagogical Implications of the Study

The findings of the study hold significant pedagogical implications for English language teaching in Turkish and European contexts, particularly regarding assessment knowledge among teachers. To start, targeted professional development programs are essential for foreign language teachers in Türkiye to enhance their language assessment knowledge. Both Turkish and European EFL teachers displayed varying levels of skill-based language assessment knowledge. Offering comprehensive training can improve teachers' grasp and application of assessment principles, ensuring more accurate and balanced language development in classrooms.

Additionally, the notable difference in language assessment knowledge between Turkish and European teachers underscores the importance of emphasizing assessment knowledge in Turkish language teacher education programs. Integrating assessment-focused courses and workshops can bridge this gap, equipping Turkish teachers with the necessary skills to assess language proficiency across all skill areas. Promoting a culture of ongoing professional development and lifelong learning can further enhance assessment knowledge among Turkish teachers. Aligning Turkish teachers' views on grades as representations of knowledge with those of European teachers can positively impact assessment practices. Providing guidance and training on interpreting and utilizing grades as indicators of language proficiency can lead to more meaningful assessment outcomes and a more student-centered approach.

Addressing these pedagogical implications can support English language teaching, fostering more effective and equitable language assessment practices that support comprehensive language development and student success.

Declaration of Conflicting Interests and Ethics

The authors declare no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in IJATE belongs to the authors. **Ethics Committee Number:** Süleyman Demirel University, 2022-6/126.14.

Contribution of Authors

Samet Fındıklı: Investigation, Resources, Visualization, Software, Formal Analysis, and Writing-original draft. **Kağan Büyükkarcı:** Methodology, Supervision, and Validation.

Orcid

Samet Fındıklı https://orcid.org/0000-0002-3477-7980

Kağan Büyükkarcı https://orcid.org/0000-0002-7365-0210

REFERENCES

Altan, M.Z. (2017). Globalization, English language teaching & Türkiye. *International Journal of Languages' Education*, *5*(4), 764–776. https://doi.org/10.18298/ijlet.2238

Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Bachman, L.F. (2004). *Statistical analyses for language assessment book (Cambridge Language Assessment)*. Cambridge University Press.

Bonnet, G. (2007). The CEFR and education policies in Europe. *The Modern Language Journal*, *91*(4), 669–672. https://doi.org/10.1111/j.1540-4781.2007.00627_7.x

Boyles, P. (2005). Assessment literacy. In M. Rosenbusch (Ed.), *National assessment summit papers* (pp. 18–23). Iowa State University.

Brown, H.D. (2000). *Principles of language learning and teaching*. Longman.

Büyükkarcı, K. (2014). Assessment beliefs and practices of language teachers in primary education. *International Journal of Instruction, 7*(1), 107-120.

Büyükkarcı, K. (2016). Identifying the areas for English language teacher development: A study of assessment literacy. *Pegem Journal of Education and Instruction, 6*(3), 333–346. https://doi.org/10.14527/pegegog.2016.017

Cheng, L., & Fox, J. (2017). *Assessment in the language classroom: Teachers supporting student learning (Applied linguistics for the language classroom)* (1st ed. 2017). Springer.

Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, *21*(3), 360-389. https://doi.org/10.1191/0265532204lt288oa

Clapham, C. (2000). Assessment and testing. *Annual Review of Applied Linguistics*, *20*, 147–161. https://doi.org/10.1017/s0267190500200093

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education*. Routledge.

Colby, D.C. (2010). *Using "Assessment for learning" practices with pre-university level ESL students: A mixed methods study of teacher and student performance and beliefs* [PhD Dissertation]. McGill University, Montreal.

Coombe, C. (2018). *An A to Z of second language assessment: How language teachers understand assessment concepts.* British Council.

Cowan, J. (1998). *On becoming an innovative university teacher.* RHE & Open University Press.

Creswell, J.W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications.

Creswell, J.W., & Clark, V.L.P. (2007). *Designing and conducting mixed methods research*. SAGE Publications.

Çakır, N. (2020). *A comparative analysis of teachers' beliefs and practices on the assessment of 4th grade-EFL students in Türkiye, Italy and Finland* [MA thesis]. Uludağ University, Bursa.

Çetin-Argün, B. (2020). *Language assessment knowledge of preservice teachers of English as a foreign language* [MA Thesis]. Çağ University, Mersin.

Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing, 25*(3), 327–347. https://doi.org/10.1177/0265532208090156

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

DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2015). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, *28*(3), 251–272. https://doi.org/10.1007/s11092-015-9233-6

Douglas, D. (2009). *Understanding language testing* (1st ed.). Routledge.

EF EPI (2022). *EF English Proficiency Index. The world's largest ranking of countries and regions by English skills.* https://www.ef.com/wwen/epi/

Fox, J. (2016). Using portfolios for assessment/alternative assessment. In E. Shohamy, I. Or, & S. May (Eds.), *Language testing and assessment. Encyclopedia of language and education* (3rd ed., pp. 135-147). Springer. https://doi.org/10.1007/978-3-319- 02261-1_9

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

Genç, E., Çalışkan, H., & Yüksel, D. (2020). Language assessment literacy level of EFL teachers: A focus on writing and speaking assessment knowledge of the teachers. *Sakarya University Journal of Education, 10*(2), 274-291. https://doi.org/10.19126/suje.626156

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. (2021). A reflection on initiatives for teachers' professional development through language assessment literacy. *Profile: Issues in Teachers' Professional Development, 23*(1), 197–213. https://doi.org/10.15446/profile.v23n1.83094

Gkogkou, E., & Kofou, I. (2021). A toolkit for the investigation of Greek EFL teachers' assessment literacy. *Languages*, *6*(188), 1-27. https://doi.org/10.3390/languages6040188

Gotch, C.M., & French, B.F. (2014). A systematic review of assessment literacy measures. *Educational Measurement: Issues and Practice*, *33*(2), 14-18. https://doi.org/10.1111/emip.12030

Green, A. (2013). *Exploring language assessment and testing: Language in action*. Routledge.

Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. *Handbook of Second Language Assessment*, 413-428. https://doi.org/10.1515/9781614513827-027

Hatipoğlu, Ç. (2010). Summative evaluation of an English language testing and evaluation course for future English language teachers in Türkiye. *English Language Teacher Education and Development (ELTED), 13*, 40-51. http://www.elted.net/uploads/7/3/1/6/7316005/v13_5hatipoglu.pdf

Hatipoğlu, Ç. (2015). English language testing and evaluation (ELTE) training in Türkiye: expectations and needs of pre-service English language teachers. *ELT Research Journal*, *4*(2), 111–128. https://dergipark.org.tr/en/download/article-file/296302

Heaton, J.B. (1989). *Writing English language tests (Longman handbooks for language teachers)*. Longman Pub Group.

Herrera, L. & Macías, 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

Hughes, A. (2003). *Testing for language teachers (2nd ed.).* Cambridge University Press.

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, L. Or, & S. May (Eds.), *Language testing and assessment* (pp. 257-270). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-02261-1_19

Jannati, S. (2015). ELT teachers' language assessment literacy: Perceptions and practices. *The International Journal of Research in Teacher Education, 6*(2), 26-37.

Jones, N.D., & Saville, N. (2009). European language policy: Assessment, learning, and the CEFR. *Annual Review of Applied Linguistics, 29*, 51-63. https://doi.org/10.1017/s0267190509090059

Kaya, T., & Mede, E. (2015). Exploring language assessment literacy of EFL instructors in language preparatory programs. *İstanbul Aydın University Journal of Education Faculty*, *7*(1), 163–189. https://doi.org/10.17932/iau.efd.2015.013/efd_v07i008

Kırkgöz, Y., Babanoğlu, M.P., & Ağçam, R. (2018). Turkish EFL teachers' perceptions and practices of foreign language assessment in primary education. *Journal of Education and E-learning Research*, *4*(4), 163–170.

Kim, A.A. (2014). Examining how teachers' beliefs about communicative language teaching affect their instructional and assessment practices: A qualitative study of EFL University instructors in Colombia. *RELC Journal, 45*(3), 337-354. https://doi.org/10.1177/0033688214555396

Krajka, J. (2019). L1 use in language tests: Investigating cross-cultural dimensions of language assessment. *Journal of Intercultural Management*, *11*(2), 107-133. https://doi.org/10.2478/joim-2019-0011

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

Lor, P. (2019). *International and comparative librarianship: Concepts and methods for global studies*. De Gruyter Saur. https://doi.org/10.1515/9783110267990

Malone, M. (2011). Assessment literacy for language educators. *CALDigest,* October, pp. 1-2.

Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi*, *168*(1), 43–60. https://doi.org/10.1501/dilder_0000000237

Mertler, C.A. (2003). *Pre-service versus in-service teachers' assessment literacy: Does classroom experience make a difference?* [Paper Presentation]. Annual Meeting of the Mid-Western Educational Research Association, Columbus, OH, USA.

Mertler, C.A., & Campbell, C. (2005). *Measuring teachers' knowledge & application of classroom assessment concepts: Development of the "Assessment Literacy Inventory"* [Paper Presentation]. Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Mohan, R. (2022). *Measurement, evaluation and assessment in education (1st ed.).* PHI Learning PVT. LTD.

Nagai, N., Birch, G.C., Bower, J.V., & Schmidt, M.G. (2020). *CEFR-informed learning, teaching and* assessment: A practical guide (Springer Texts in Education) (1st ed. 2020 ed.). Springer.

O'Loughlin, K. (2006). Learning about second language assessment: Insights from a postgraduate student on-line subject forum. *University of Sydney Papers in TESOL*, 1, 71–85.

Ölmezer-Öztürk, E., & Aydin, B. (2018). Toward measuring language teachers' assessment knowledge: Development and validation of Language Assessment Knowledge Scale (LAKS). *Language Testing in Asia*, 8(1). https://doi.org/10.1186/s40468-018-0075-2

Ölmezer-Öztürk, E., & Aydın, B. (2019). Voices of EFL teachers as assessors: Their opinions and needs regarding language assessment. *Journal of Qualitative Research in Education, 7*(1), 373-390. doi:10.14689/issn.21482624.1.7c1s.17m

Popham, W.J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator, 46*(4), 265-273.

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

Sarıyıldız, G. (2018). *A study into language assessment literacy of preservice English as a foreign language teachers in Turkish context* [MA thesis]. Hacettepe University, Ankara.

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

Sevimel-Şahin, A. (2019). *Exploring foreign language assessment literacy of pre-service English language teachers* [Doctoral dissertation]. Anadolu University, Eskişehir, Türkiye.

Sevimel-Sahin, A., & Subaşı, G. (2021). Exploring foreign language assessment literacy training needs of pre-service English language teachers. *International Online Journal of Education and Teaching (IOJET), 8*(4), 2783-2802.

Shepard, L.A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14.

Siegel, M.A., & Wissehr, C. (2011). Preparing form the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education, 22*(4), 371-391. https://doi.org/10.1007/978-3-319-02261-1_19

Stiggins, R. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan, 77*(3), 238–245.

Şahin, S. (2009). *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* [Master's Thesis]. Middle East Technical University.

Tamerer, R.B. (2019). *An investigation of Turkish pre-service EFL teachers' language assessment literacy* [MA thesis]. Kocaeli University, Kocaeli.

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

Thomas, J., Allman, C., & Beech, M. (2004). Assessment for the diverse classroom: A handbook for teachers. *Tallahassee, FL: Florida Department of Education, Bureau of Exceptional Education and Student Services.*

Turkish Ministry of Education. (2018). *High-school English language teaching program.* [Curriculum]. https://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=342

Ünlücan-Tosun, F., & Glover, P. (2020). How do school teachers in Turkey perceive and use the CEFR? *International Online Journal of Education and Teaching (IOJET), 7*(4), 1731-1739.

Vogt, K., Guerin, E., Sahinkaraks, S., Pavlou, P., Tsagari, D., & Afiri, Q. (2008). *Assessment literacy of foreign language teachers in Europe.* Poster session presented at the 5th Annual meeting of EALTA, Athens, Greece.

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

Wardani, W.O., Sukyadi, D., & Purnawarman, P. (2021). Exploring teacher assessment literacy in EFL classroom. *Proceedings of the 2nd International Conference on Progressive*

Education, ICOPE 2020, 16–17 October 2020, Universitas Lampung, Bandar Lampung, Indonesia. https://doi.org/10.4108/eai.16-10-2020.2305253

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37*(1), 3–14. https://doi.org/10.1016/j.stueduc.2011.03.001

Xu, Y., & Brown, G.T.L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, *58*(1), 149-162. https://doi.org/10.1016/j.tate.2016.05.010

Xu, Y., & Brown, G.T.L. (2017). University English teacher assessment literacy: A survey-test report from China. *Papers in Language Testing and Assessment, 6*(1), 133–158.

Yetkin, C. (2015). *An investigation on ELT teacher candidates' assessment literacy* [MA thesis]. Cag University, Mersin.