



## **Language Assessment Practices and Training Preferences of EL Teachers: Iranian EFL Teachers in Focus**

**Kaveh Jalilzadeh** (Corresponding Author)

*Foreign Languages Department, Istanbul University, Turkey*

[Kaveh.j@iuc.edu.tr](mailto:Kaveh.j@iuc.edu.tr)

**Adel Dastgoshadeh**

*English Language and Literature Department, Sanandaj Branch, Islamic Azad University, Iran*

[Adastgoshadeh@gmail.com](mailto:Adastgoshadeh@gmail.com)

**Raheleh Khosravi**

*English Language and Literature Department, Kurdistan University, Iran*

[rahelekhosravi52@gmail.com](mailto:rahelekhosravi52@gmail.com)

### **ARTICLE INFO:**

**Received date:**

**2023.11.17**

**Accepted date:**

**2023.11.30**

**Print ISSN: 2251-7995**

**Online ISSN: 2676-6876**

### **Keywords:**

language assessment;  
assessment practices; training  
preferences; assessment  
literacy.



### **Abstract**

This research explores language assessment practices and training preferences in Iranian English as a Foreign Language (EFL) teaching, aiming to provide valuable insights into the current landscape among 363 Iranian EFL teachers. Data collection included diverse demographics, facilitating a thorough analysis of assessment practices and preferences. Statistical analyses, such as chi-square tests, revealed a significant gap between the perceived importance and the actual proficiency of Iranian EFL teachers. Speaking skills are prioritized while listening comprehension is least emphasized. Common assessment methods include active class participation, oral presentations, and closed-ended tests, with underutilized methods suggesting a need for broader teacher development programs. The study underscores the diverse terminology used for teacher-mediated assessments, emphasizing the multifaceted nature of EFL assessment practices. In summary, the paper highlights the significance of tailored assessment literacy programs to bridge the gap and enhance English language teaching in Iran.

DOI: 10.22034/elt.2023.59200.2583

Citation: Jalilzadeh, K., Dastgoshadeh, A., & Khosravi, R., (2023). Language Assessment Practices and Training Preferences of EL Teachers: Iranian EFL Teachers in Focus. *Journal of English Language Teaching and Learning*, 15(32), 144-160. DOI: 10.22034/elt.2023.59200.2583

## Introduction

Language assessment plays a multifaceted and pivotal role in the realm of English Language Teaching (ELT), serving as a means to gauge students' learning progress, identify their strengths and weaknesses, and inform instructional practices (Brown, 2004; Harlen, 2012). Within the context of English as a Foreign Language (EFL) instruction in Iran, this process assumes a critical role in shaping the educational landscape. However, empirical evidence underscores a significant disparity between the importance of language assessment and the proficiency of EFL teachers, particularly in the Iranian context (Naraghizadeh, Azizmalayeri, & Khalaji, 2022; Abbasian, & Khadempir, 2018). This disconnect gives rise to several challenges, including the risk of inaccurately assessing students' learning, the improper utilization of assessment tools, and the potential demotivation of students (Brown, 2004).

Several factors contribute to the deficiency in language assessment literacy among Iranian EFL educators. Firstly, the absence of comprehensive language assessment training within the existing teacher education programs in Iran impedes the development of proficient assessors (Babamoradi, Nassiri & Ahmadi, 2018). Secondly, the prevalence of high-stakes standardized testing in the Iranian educational system inadvertently steers teachers toward teaching the test, rather than fostering students' language proficiency (Afshari & Ghafar, 2018; Harlen, 2012). Additionally, limited access to high-quality language assessment resources and materials exacerbates these challenges for Iranian EFL teachers.

In response to these assessment-related challenges, there has been a global shift from traditional, summative assessment practices, often referred to as "assessment of learning" (AOL), to a more dynamic and formative approach known as "assessment for learning" (AFL) (Assessment Reform Group, 2002; Black and Wiliam, 1998; Black et al., 2003; Stiggins, 2005, 2008). AFL signifies a transformation in the assessment culture, where assessment activities are intimately intertwined with classroom instruction, primarily aimed at enhancing student learning (Black & Wiliam, 1998). It encourages continuous monitoring and scaffolding of students, aligning assessment with pedagogical objectives, and granting learners a greater degree of autonomy in their educational journey (Black & Wiliam, 1998; Pat-El et al., 2013).

While AFL has gained prominence in educational discourse, the field of English language teaching encompasses a range of terms and concepts, often used interchangeably, such as formative assessment, teacher-based assessment, classroom-based assessment, and alternative assessment. These terms underscore the importance of teacher-mediated, context-based, and classroom-embedded assessments, which stand in contrast to traditional external examinations (Davison & Leung, 2009). In the context of this paper, AFL encompasses this diverse array of practices, emphasizing not only formative assessment but also the broader shift in assessment culture (Black & Wiliam, 1998). This research aims to investigate the language assessment practices and training preferences of Iranian EFL teachers, shedding light on the challenges they face and the opportunities for enhancing their assessment competencies. As AFL gains momentum, particularly in EFL settings, it becomes imperative to understand its implementation within the Iranian educational landscape. Drawing upon a rich tapestry of

research studies, government reports, and insights from EFL teachers, this paper seeks to contribute to the ongoing discourse on language assessment in the Iranian context.

In order to effectively address the gaps in existing literature and contribute to the advancement of knowledge in this field, the following research questions will guide our investigation:

1. Which skills are more emphasized in the assessment of English language teaching in the Iranian context?
2. How do Iranian English language teachers (ELTs) assess their students?
3. What type of feedback do Iranian ELTs provide for their learners as a result of their assessment?
4. What are the preferred formats for a training event offered in an online learning environment on language assessment?
5. What is the effectiveness of the provided online training materials in terms of their practical utility and impact on skill development?

## Literature

### Assessment Literacy

The concept of assessment literacy has its origins in general education but has since gained prominence in the field of language education, particularly concerning assessment and measurement practices in educational contexts. Over time, assessment literacy has become a subject of increased research interest, with scholars such as Fulcher (2012), Taylor (2009), and Walters (2010) noting that there is no universally agreed-upon definition for this concept.

The term "assessment literacy" was first introduced by Stiggins (1991a, 1995) within the context of general education. It was used to describe the idea that classroom teachers should be able to distinguish between effective and ineffective assessment practices. Stiggins (1999) further elaborated that "assessment literacy" has become a common term referring to the skills and knowledge necessary for stakeholders to navigate the evolving landscape of assessment. Despite various recommended approaches aimed at fostering the development of assessment literacy, a consensus on its exact components remains elusive. Inbar-Lourie (2008) extended the concept of assessment literacy to include teachers' ability to recognize the social dimensions of assessment and the connections between language knowledge and assessment task types. Moreover, Language Assessment Literacy (LAL) was defined as encompassing the knowledge, skills, and principles related to test construction, interpretation, utilization, and the cultivation of a critical perspective on assessment's role within the broader educational context (O'Loughlin, 2013). It is widely acknowledged that teachers need to assess students' progress, but many lack a solid foundation in assessment fundamentals (Popham, 2009). LAL competence involves understanding both the "what" (construct) and the "how" (method) of language testing and assessment (Shohamy, 2001).

Inbar-Lourie (2008) emphasized that language assessment knowledge is not a singular concept but rather a fusion of assessment literacy skills and language-specific competencies, forming a distinct entity known as "language assessment literacy" (p. 389). Similarly, Malone (2013) referred to assessment literacy as the degree of familiarity that language educators have

with terminologies in language testing and their ability to apply this knowledge to classroom practices, especially concerning language assessment. Popham (2009) stressed the significance of assessment literacy for teachers' professional growth, considering it an essential objective for their development. McMillan (2014), on the other hand, pointed out that traditional or objective assessments primarily gauge lower-order thinking abilities, while newer alternatives in language assessment focus on evaluating higher-order thinking abilities. This alternative assessment, often referred to as authentic assessment, pertains to tasks that relate to students' everyday lives, reflecting meaningful and appropriate achievements (Bachman and Palmer, 1996).

A central component of teachers' classroom assessment literacy is their knowledge base related to assessment. According to Bandura, Freeman, and Lightsey (1999) and Fishbein and Ajzen (2011), individual teachers' understanding and expertise in this area are crucial factors influencing the effectiveness of assessment practices. Farhady and Tavassoli (2018) underscored the significance of Language Assessment Knowledge (LAK) for EFL teachers, as it aids in creating appropriate assessments, making informed decisions about student development and performance, and ultimately enhancing their professional achievements. Furthermore, Boud and Falchicov (2006), Joughin (2008), and Earl (2013) have observed a strong connection between the quality of assessment and instruction and student achievements.

As a result, it is essential for language educators to have a comprehensive understanding of assessment and use it effectively in their teaching contexts. Experts in the field have argued that the depth of teachers' assessment knowledge significantly impacts their ability to implement assessments that improve instruction and promote student learning (Stiggins, 1991a, 1995; Popham, 2006, 2009). Price et al. (2012) contend that language educators need comprehensive knowledge of language assessment to navigate the complexities of assessment in the classroom. This knowledge is instrumental in understanding the entire assessment process, making informed decisions about what skills and abilities to measure, and selecting and applying assessment tasks to evaluate student performance.

Overall, assessment literacy empowers teachers to make well-informed decisions about assessment practices, leading to improved student learning outcomes. It involves the ability to select appropriate assessments, interpret and utilize assessment results effectively, and apply this knowledge to enhance instructional strategies and educational programs. By possessing assessment literacy, teachers can ensure that assessment becomes a valuable tool for fostering student growth and success in the classroom (Kahl, Hofman, & Bryant, 2013; Jalilzadeh et al., 2023).

In addition to their expertise in content, language, and EFL instruction, EFL teachers must also possess assessment literacy, a critical facet that provides them with the knowledge and tools required to comprehend what they are evaluating, how to tailor assessments to specific purposes, and the decisions necessary for effective learner assessment and optimized learning outcomes (Djoub, 2017). The acquisition of assessment literacy, as emphasized by Coombe et al. (2009, cited in Djoub, 2017), is contingent upon the teacher's clear grasp of the meaning of effective assessment within the educational context, shaping their approach to assessment, and exposure to effective training, whether through online platforms or hands-on workshops.

Additionally, the abundance of diverse assessment resources and a willingness to embrace educational change play a pivotal role in fostering assessment literacy.

### **Assessment Training**

Teachers who have undergone comprehensive training in CA are more likely to embrace formative assessment strategies that actively engage students in the learning process (Black & William, 1998). They may prefer methods that allow for ongoing data collection, such as classroom observations, peer assessment, and self-assessment, as these techniques align with their training and the principles of CA for learning improvement (Earl & Katz, 2006).

Additionally, training programs that emphasize the importance of aligning assessment with learning objectives and curriculum standards can influence teachers' preferences in designing assessments (Wiggins & McTighe, 1998). Educators who have received training in criterion-referenced assessment may prioritize the development of clear and specific learning outcomes and assessments that directly measure these outcomes (Gronlund, 1998).

Moreover, teacher training practices can introduce educators to innovative assessment tools and technologies (Bennett, 2011). Teachers who have been exposed to digital assessment platforms or e-portfolios may be more inclined to incorporate these modern tools into their CA practices (Pellegrino, Chudowsky, & Glaser, 2001). These preferences can have a profound impact on the efficiency and effectiveness of assessment in EFL classrooms.

Teacher training practices play a crucial role in shaping educators' preferences and competence in conducting effective classroom assessment. In Iranian EFL education, as in many other educational systems, teachers' ability to implement CA effectively is often influenced by the quality and nature of their professional development and training. Teacher training programs that focus on CA can significantly impact teachers' understanding of assessment principles, their awareness of various assessment methods, and their ability to align assessment practices with educational goals (Stiggins, 2002). In essence, teacher training practices and preferences in CA are intertwined. Effective training programs not only equip educators with the knowledge and skills needed for sound assessment practices but also influence their attitudes and preferences. As such, exploring the relationship between teacher training experiences and CA practices in the Iranian EFL context can provide valuable insights into ways to enhance the quality of assessment in language education.

## **Methodology**

### **Research Design and Data Analysis**

This study aimed to engage a broad spectrum of Iranian EFL (English as a Foreign Language) teachers, striving to ensure comprehensive insights into the assessment practices and training preferences within this particular educational context. To accomplish this, the primary data collection method employed was a questionnaire survey—a methodology traditionally favored within the field for examining the language assessment needs of distinct stakeholder groups, such as policymakers (Pill & Harding, 2013), fellow educators (Fulcher, 2012; Vogt, Tsagari, & Csépes, 2020; Xu & Brown, 2016), students (Vogt, Tsagari, & Spanoudis, 2020), and university admissions advisors (Deygers & Malone, 2019).

Despite relying on a single data collection technique, this research endeavor ensured data richness by encompassing both pre-service and in-service English teachers operating at primary and secondary education levels, in line with the holistic approach advocated by Nunan and Bailey (2009).

The survey instrument used was adapted from Vogt, Sperling, and Brüstle (2018), a tool that had undergone rigorous piloting and been deployed across diverse international contexts, including Greece, Cyprus, Hungary, and Germany, as part of the Needs Analysis of the Erasmus+ TALE project ([www.taleproject.eu](http://www.taleproject.eu)). The questionnaire's questions and scope were thoughtfully tailored to align with the specific objectives of the present study.

Comprehensively addressing a wide array of assessment-related knowledge and skills, the questionnaire was designed in accordance with established models articulated by Davies (2008), Inbar-Lourie (2008), and Fulcher (2012). The questionnaire was structured into four main sections: I. General Information, II. Assessment Practices, III. Assessment Profiles and Training Needs, and IV. Use of Technology. Each of these sections further delved into subcategories, comprising both closed and open-ended questions. For data analysis, descriptive and inferential statistics, including percentages and means, and chi-square were employed, and a detailed presentation of the findings will be provided in the subsequent sections of this study.

### **Participants**

In this study, a total of 363 Iranian EFL (English as a Foreign Language) teachers from private institutes participated, comprising 243 females and 120 males, with ages spanning from 23 to over 56. Their educational backgrounds encompassed a wide range, including associate, bachelor's, master's, and even doctorate degrees, alongside individuals without formal degrees who were self-learners or practitioners. Their teaching experience ranged from pre-service educators to those with over 15 years of experience, offering a diverse perspective. Furthermore, the age range of their learners varied from 6 to over 18 years, allowing for a comprehensive exploration of assessment practices across different age groups. To uphold ethical research standards, the study adhered to the Guidelines for Research Ethics in the Social Sciences, Humanities, Law, and Theology established by The Norwegian National Research Ethics Committee in 2019. Participation in the survey was entirely voluntary, with measures in place to ensure that responding to the questionnaire, which could be completed comfortably within 15-20 minutes, did not overburden the participants. The distribution of the survey questionnaire was facilitated through email communication with individuals and institutions, as well as via social media platforms such as Facebook. This distribution was complemented by a succinct invitation detailing the study's objectives and a consent form to maintain ethical integrity.

**Table 1.** Demographic Information of the participants

Variable	Categories	Frequency	Percent	
Gender	Female	243	66.9	
	Male	120	33.1	
Age	Under 25	118	32.5	
	26-35	101	27.8	
	36-45	95	26.2	
	46-55	42	11.6	
	56+	7	1.9	
	Degree	Associate degree	30	8.3
	Bachelor degree	135	37.2	
	Master degree	119	32.8	
	Philosophy doctor	53	14.6	
	None of them	19	5.2	
Experience	1-5	98	27.0	
	6-10	58	16.0	
	10-15	58	16.0	
	15+	73	20.1	
	pre-service teacher	76	20.9	
	6-12	68	18.7	
	13-15	40	11.0	
	13-15, 16-17	15	4.1	
	13-15, 16-17, over 18	19	5.2	
	6-12, 13-15	13	3.6	
	6-12, 13-15, 16-17	11	3.0	
	6-12, 13-15, 16-17, over 18	25	6.9	
	13-15, over 18	1	.3	
	16-17	32	8.8	
	16-17, over 18	24	6.6	
	Learners' Age Range	6-12, 16-17	2	.6
		6-12, 16-17, over 18	1	.3
6-12, over 18		7	1.9	
over 18		105	28.9	
Do you need training?	No	107	29.5	
	Yes	256	70.5	

### Data Analysis

The analysis of the collected data in this study was conducted systematically to derive meaningful insights into the language assessment practices and training preferences of Iranian EFL teachers.

#### 1. Which skills are more emphasized in the assessment of English language teaching in the Iranian context?

Table 2 displays the frequencies and percentages for the areas emphasized in the assessment of English language. The results showed that speaking ( $n = 272$ , 18.4 %) was the most emphasized area in English language assessment. This was followed by reading ( $n = 257$ , 17.4 %) and grammar ( $n = 254$ , 17.2 %). The least emphasized area was listening comprehension ( $n = 222$ , 15 %).

**Table 2.** Frequencies and Percentages of Areas Emphasized in English Language Assessment

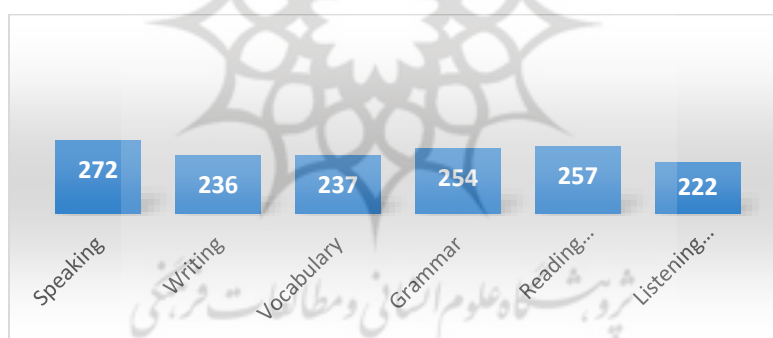
	Frequency	Percent	Residual
Speaking	272	18.4	25.7
Writing	236	16.0	-10.3
Vocabulary	237	16.0	-9.3
Grammar	254	17.2	7.7
Reading Comprehension	257	17.4	10.7
Listening Comprehension	222	15.0	-24.3
Total	1478	100.0	

Table 3 displays the results of chi-square. The results ( $\chi^2(5) = 6.56$ ,  $p > .05$ , Cramer's  $V = .066$  representing a weak effect size) indicated that there were no significant differences between the frequencies observed in Table 4.1.

**Table 3.** Chi-Square Statistics; Areas Emphasized in English Language Assessment

	Comments
Chi-Square	6.566 <sup>a</sup>
Df	5
Asymp. Sig.	.255
Cramer's V	.095

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 246.3.

**Figure 1.** Frequencies of areas emphasized in English language assessment

## 2. How do Iranian English language teachers (ELTs) assess their students?

As displayed in Table 4.3, active class participation (49.1 %) was the most frequent method of English language assessment. This was followed by an oral presentation (39.5 %) and a test with closed-ended answers (33 %). The least frequent methods were translation (17.8 %) and portfolio assessment (16.8 %).

Any standardized residuals higher than 1.96 indicate that the observed frequency was significantly higher than what was expected. On the other hand; any standardized residuals higher than -1.96 indicate that the observed frequency was significantly lower than what was expected. Thus, it can be concluded that the employment of the following assessment methods was significantly beyond what was expected; oral presentation (std. residual = 5 > 1.96), closed-ended tests (std. residual = 2.6 > 1.96), and class participation (std. residual = 8.6 > 1.96). Following the same logic, it can be concluded that the following methods were employed



significantly below what was expected; open-ended tests (std. residual = -2.4 > -1.96), portfolio assessment (std. residual = -3.9 > -1.96), peer assessment (std. residual = -3.7 > -1.96), self-assessment (std. residual = -2.8 > -1.96), and translation (std. residual = -2.7 > -1.96).

**Table 4.** *Frequencies, Percentages and Standardized Residuals: Methods of Assessment*

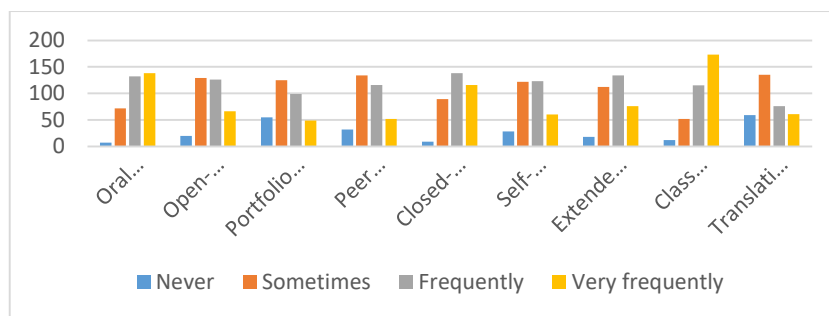
		Choices				Total
		Never	Sometimes	Frequently	Very frequently	
Oral Presentation	Count	7	72	132	138	349
	%	2.0%	20.6%	37.8%	39.5%	100.0%
	Std. Residual	-3.9	-3.7	1.0	5.0	
Open-ended Tests	Count	20	129	126	66	341
	%	5.9%	37.8%	37.0%	19.4%	100.0%
	Std. Residual	-1.3	2.0	.7	-2.4	
Portfolio Assessment	Count	55	125	99	49	328
	%	16.8%	38.1%	30.2%	14.9%	100.0%
	Std. Residual	5.8	2.1	-1.4	-3.9	
Peer Assessment	Count	32	134	116	52	334
	%	9.6%	40.1%	34.7%	15.6%	100.0%
	Std. Residual	1.1	2.7	.0	-3.7	
Closed-ended Test	Count	9	89	138	116	352
	%	2.6%	25.3%	39.2%	33.0%	100.0%
	Std. Residual	-3.5	-2.1	1.5	2.6	
Self-Assessment	Count	28	122	123	60	333
	%	8.4%	36.6%	36.9%	18.0%	100.0%
	Std. Residual	.4	1.6	.7	-2.8	
Extended Writing	Count	18	112	134	76	340
	%	5.3%	32.9%	39.4%	22.4%	100.0%
	Std. Residual	-1.7	.4	1.5	-1.3	
Class Participation	Count	12	52	115	173	352
	%	3.4%	14.8%	32.7%	49.1%	100.0%
	Std. Residual	-3.0	-5.6	-.6	8.6	
Translation	Count	59	135	76	61	331
	%	17.8%	40.8%	23.0%	18.4%	100.0%
	Std. Residual	6.5	2.9	-3.6	-2.7	
Count		240	970	1059	791	3060
%		7.8%	31.7%	34.6%	25.8%	100.0%

Table 4 displays the results of chi-square test ( $\chi^2(24) = 373.53$ ,  $p < .05$ , Cramer's  $V = .202$  representing a weak effect size) indicated that there were significant differences between the frequencies observed in Table 4.2. The results of standardized residuals were discussed above.

**Table 5.** *Chi-Square Tests; Methods of Assessment*

	Value	Df	Asymptotic Significance (2-sided)
Pearson Chi-Square	373.535 <sup>a</sup>	24	.000
Likelihood Ratio	365.357	24	.000
Linear-by-Linear Association	.884	1	.347
N of Valid Cases	3060		
Cramer's V	.202		.000

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 25.73.



**Figure 2.** Frequencies of methods English language assessment

### 3. What type of feedback do Iranian ELTs provide for their learners as a result of their assessment?

As displayed in Table 4.5, putting a mark ( $n = 219$ , 28 %) was the most frequent type of feedback. This was followed by giving brief comments ( $n = 201$ , 25.7%), hints and comments ( $n = 194$ , 24.8 %), and giving detailed comments ( $n = 167$ , 21.4 %).

**Table 6.** Frequencies and Percentages; Types of Feedback

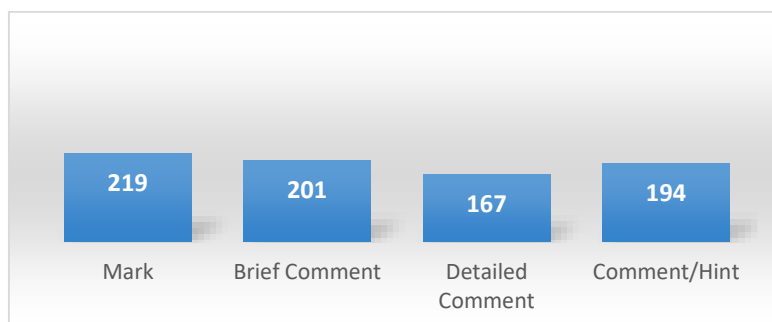
	Frequency	Percent	Residual
Mark	219	28.0	23.8
Brief Comment	201	25.7	5.8
Detailed Comment	167	21.4	-28.2
Comment/ Hint	194	24.8	-1.2
Total	781	100.0	

The results of chi-square ( $\chi^2(3) = 7.154$ ,  $p > .05$ , Cramer's  $V = .095$  representing a weak effect size) indicated that there were no significant differences between the frequencies observed in Table 4.5. The results of standardized residuals were discussed above.

**Table 7.** Chi-Square Statistics; Types of Feedback

	Comments
Chi-Square	7.154 <sup>a</sup>
Df	3
Asymp. Sig.	.067
Cramer's V	.095

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 195.3.



**Figure 3.** Frequencies of methods English language assessment

#### 4. What are the preferred formats for a training event offered in an online learning environment on language assessment?

As displayed in table 8 combination of face to face and online courses (60.9 %) was the most preferred format of training. This was followed by interactive online sources (54 %), online resources (43.7 %), and printed materials (41.2 %).

The std. residuals indicated that the frequency for the combination of face to face and online was significantly beyond what was expected (Std. Residual = 2.9 > 1.96). The frequency for printed materials was significantly below expectation (Std. Residual = -2.3 > -1.96)

**Table 8.** Frequencies, Percentages and Standardized Residuals; Preferred Formats of Training

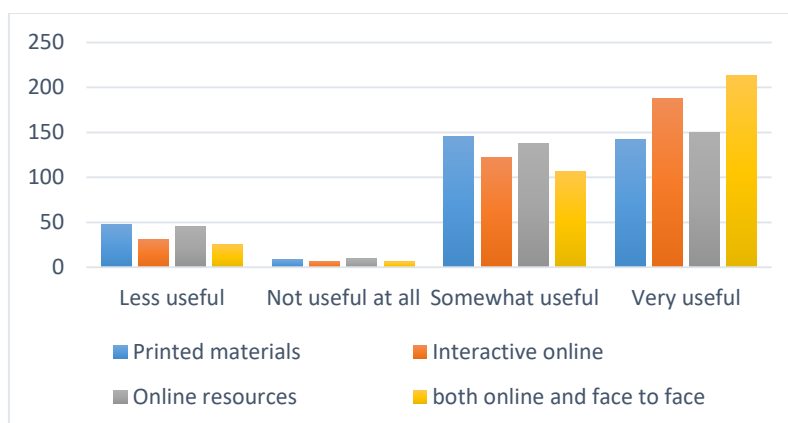
		Preferred Format of Training				Total
		Less useful	Not useful at all	Somewh at useful	Very useful	
Printed Materials	Count	48	9	146	142	345
	%	13.9%	2.6%	42.3%	41.2%	100.0%
	Std. Residual	1.8	.4	1.6	-2.3	
Interactive Online courses	Count	31	7	122	188	348
	%	8.9%	2.0%	35.1%	54.0%	100.0%
	Std. Residual	-1.0	-.4	-.6	1.1	
Online Resources	Count	45	10	138	150	343
	%	13.1%	2.9%	40.2%	43.7%	100.0%
	Std. Residual	1.3	.7	1.0	-1.6	
Face to face + Online	Count	25	6	106	213	350
	%	7.1%	1.7%	30.3%	60.9%	100.0%
	Std. Residual	-2.1	-.7	-2.0	2.9	
Total	Count	149	32	512	693	1386
	%	10.8%	2.3%	36.9%	50.0%	100.0%

The results of chi-square ( $\chi^2 (9) = 37.39$ ,  $p < .05$ , Cramer's  $V = .095$  representing a weak effect size) indicated that there were significant but weak differences between the frequencies observed in Table 8. The results of standardized residuals were discussed above.

**Table 9.** Chi-Square Tests; Preferred Formats of Training

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	37.391 <sup>a</sup>	9	.000
Likelihood Ratio	37.684	9	.000
N of Valid Cases	1386		
Cramer's V	.095		.000

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.92.



**Figure 4.** Frequencies of preferred formats of training

### 5. How useful are these following online training materials?

As displayed in Table 4.9 frequencies and percentages for the usefulness of online training materials. The most useful materials for online training were short videos (59 %). This was followed by a discussion with colleagues (54.3 %), practical materials and evaluating assessment materials (50.3 %), and materials to read (48.3 %). As displayed in Table 4.9, none of the frequencies were higher than  $\pm 1.96$  under the last column; i.e. very useful. That is to say, none of the training materials were identified significantly higher or lower than what was expected.

**Table 10.** Frequencies and Percentages; Usefulness of Online Training Materials

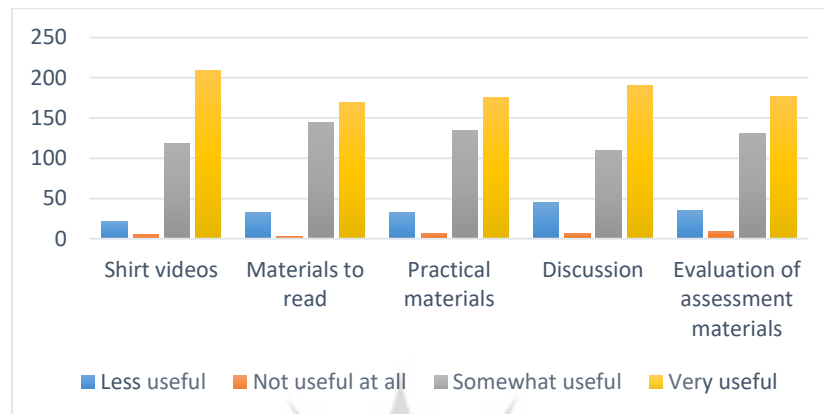
		Usefulness of Online Training Materials				Total
		Less useful	Not useful at all	Somewhat useful	Very useful	
Short Video	Count	22	5	118	209	354
	%	6.2%	1.4%	33.3%	59.0%	100.0%
	Std. Residual	-2.0	-.3	-.9	1.7	
Materials to read	Count	33	3	145	169	350
	%	9.4%	0.9%	41.4%	48.3%	100.0%
	Std. Residual	-.1	-1.2	1.6	-1.1	
Practical materials	Count	33	6	134	175	348
	%	9.5%	1.7%	38.5%	50.3%	100.0%
	Std. Residual	-.1	.1	.7	-.6	
Discussion	Count	45	6	110	191	352
	%	12.8%	1.7%	31.3%	54.3%	100.0%
	Std. Residual	2.0	.1	-1.6	.5	
Evaluation of Materials	Count	35	9	131	177	352
	%	9.9%	2.6%	37.2%	50.3%	100.0%
	Std. Residual	.2	1.3	.3	-.6	
Total	Count	168	29	638	921	1756
	%	9.6%	1.7%	36.3%	52.4%	100.0%

The results of chi-square ( $\chi^2 (12) = 22.59, p < .05$ , Cramer's  $V = .065$  representing a weak effect size) indicated that there were significant but weak differences between the frequencies observed in Table 4.9. The results of standardized residuals were discussed above.

**Table 11.** Chi-Square Tests; Usefulness of Online Training Materials

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	22.594 <sup>a</sup>	12	.031
Likelihood Ratio	22.817	12	.029
N of Valid Cases	1756		
Cramer's V	.035		.031

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 5.75.

**Figure 5.** Frequencies of preferred usefulness of online training materials

## Discussion

This study provides valuable insights into the assessment practices and training preferences of Iranian EFL teachers, shedding light on the complex landscape of language assessment in the Iranian educational context. The findings of this research not only align with prior studies but also offer critical implications for teacher education and the enhancement of English Language Teaching (ELT) in Iran.

Firstly, the study highlights a significant gap between the perceived importance of language assessment and the proficiency of Iranian EFL teachers in conducting assessments. This gap has been a recurring concern in the literature (Naraghizadeh, Azizmalayeri, & Khalaji, 2022; Abbasian, & Khadempir, 2018). The repercussions of this misalignment are profound, encompassing the accuracy of student evaluations, the proper utilization of assessment tools, and the potential for student demotivation. These issues resonate with the global discourse on assessment literacy and the urgent need for targeted training in this domain. Therefore, the findings underscore the pressing necessity for comprehensive assessment literacy programs tailored to Iranian EFL teachers. Such programs should encompass a broad spectrum of assessment methods and strategies, equipping teachers with the skills and knowledge necessary to conduct effective assessments.

Secondly, the study identifies prevalent assessment practices among Iranian EFL teachers, including active class participation, oral presentations, and closed-ended tests. These practices are in line with the global shift towards Assessment for Learning (AFL), where assessment is closely integrated with instruction to enhance student learning (Black and Wiliam, 1998). Such alignment with AFL principles is encouraging, as it promotes a learner-centered and supportive classroom environment. However, the underutilization of certain assessment methods like

open-ended tests, portfolio assessment, peer assessment, self-assessment, and translation reveals potential areas for improvement. These underused methods represent untapped opportunities to diversify assessment practices and cater to the diverse needs and learning styles of students. Consequently, this aspect of the findings highlights the need for professional development initiatives that expand the repertoire of assessment tools available to Iranian EFL teachers. These initiatives should encourage teachers to explore and incorporate a wider range of assessment methods into their instructional practices.

Furthermore, the study's observations on the diverse terminology used to describe teacher-mediated, context-based, and classroom-embedded assessments resonate with the broader discourse on the assessment culture in EFL settings (Davison & Leung, 2009). This diversity in terminology underscores the multifaceted nature of assessment practices, with various terms like formative assessment, teacher-based assessment, and alternative assessment all pointing to the role of assessments that occur within the classroom context. This recognition of the diverse assessment landscape should be leveraged to foster an inclusive and holistic approach to language assessment. It emphasizes the importance of accommodating alternative assessment practices that go beyond traditional testing methods and facilitate comprehensive language learning experiences.

In summary, this research not only reinforces the alignment of its findings with existing literature but also offers actionable insights for enhancing assessment practices and training approaches for Iranian EFL teachers. The study's emphasis on the gap between perceived importance and actual proficiency, the prevalence of certain assessment practices, and the underutilization of others, all contribute to a nuanced understanding of the assessment landscape in Iran. Moving forward, targeted efforts in teacher education and professional development should prioritize assessment literacy and the exploration of a wider array of assessment methods. These endeavors hold the potential to empower Iranian EFL teachers to provide more accurate and meaningful feedback and, in turn, enhance the overall quality of English language education in the country.

### **Conclusion**

In conclusion, this study illuminates the assessment practices and training preferences of Iranian EFL teachers, revealing a substantial gap between the recognized importance of language assessment and teachers' proficiency in conducting assessments. While prevalent practices align with global trends in Assessment for Learning (AFL), such as active class participation and oral presentations, the underutilization of diverse assessment methods underscores the need for comprehensive assessment literacy programs. The study's recognition of the diverse terminology used for teacher-mediated assessments highlights the multifaceted nature of classroom assessment in EFL settings. To enhance the quality of English Language Teaching in Iran, there is an imperative to bridge this gap through targeted professional development initiatives that empower teachers to employ a broader range of assessment tools and strategies, thereby fostering more accurate and meaningful student evaluation and facilitating holistic language learning experiences.

## References

- Abbasian, G., & Khadempir, F. (2018). Implementation and Assessment Challenges in Iranian Secondary High School EFL Program. *Journal of English Language Pedagogy and Practice*, 11(23), 1-20. DOI: 10.30495/jal.2019.664537
- Afshari, H., & Ghafar, M. A. (2018). The washback effect of high-stakes tests on EFL teachers' teaching: Perceptions and observations. *Language Testing in Asia*, 8(1), 1-18.
- Assessment Reform Group. (2002). *Assessment for learning: 10 principles*. University of Cambridge, Faculty of Education.
- Babamoradi, P., Nasiri, M., & Mohammadi, E. (2018). Learners' Attitudes toward Using Dynamic Assessment in Teaching and Assessing IELTS Writing Task One. *International Journal of Language Testing*, 8(1), 1-11.
- Bachman, L. F. (2001). Designing and developing useful language tests. Experimenting with uncertainty: *Essays in honour of Alan Davies*, 109-116.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1). Oxford University Press.
- Bandura, A., Freeman, W. H., & Lightsey, R. (1999). *Self-efficacy: The exercise of control*.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5-25.
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Assessment in Education: Principles, Policy & Practice*, 16(1), 5-31.
- Black, P., Harrison, C., & Lee, C. (2003). *Assessment for learning: Putting it into practice*. McGraw-Hill Education (UK).
- Boud, D., & Falchikov, N. (2006). Aligning assessment with long term learning. *Assessment & evaluation in higher education*, 31(4), 399-413.
- Brown, J. D. (2004). Performance assessment: Existing literature and directions for research. *Second Language Studies*, 22 (2), 91-139. Google Scholar
- Coombe, C., Troudi, S., & Al-Hamly, M. (2012). Foreign and second language teacher assessment literacy: Issues, challenges, and recommendations. *The Cambridge guide to second language assessment*, 1, 20-29.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language testing*, 25(3), 327-347.
- Davison, C., & Leung, C. (2009). Current issues in English language teacher based assessment. *TESOL quarterly*, 43(3), 393-415.
- Deygers, B., & Malone, M. E. (2019). Language assessment literacy in university admission policies, or the dialogue that isn't. *Language Testing*, 36(3), 347-368.
- Djoub, Z. (2017). Assessment Literacy: Beyond Teacher Practice. In Al-Mahrooqi et al. (Eds.), *Revisiting EFL Assessment, Second Language Learning and Teaching* (pp. 9-27), Springer International Publishing Switzerland.
- Earl, L. M. (2013). Assessment for learning; Assessment as learning: Changing practices means changing beliefs. *Assessment*, 80, 63-71.
- Earl, L. M., & Katz, S. (Eds.). (2006). *Leading Schools in a Data-Rich World: Harnessing data for school improvement*. Corwin Press.

- Farhady, H., & Tavassoli, K. (2018). Developing a language assessment knowledge test for EFL teachers: A data-driven approach. *Iranian Journal of Language Teaching Research*, 6(3) (Special Issue), 79-94.
- Fishbein, M., & Ajzen, I. (2011). *Predicting and changing behavior: The reasoned action approach*. Taylor & Francis.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132.
- Glaser, R. (1963). Instructional technology and the measurement of learning outcomes: Some questions. *American Psychologist*, 18(8), 519-521.
- Gronlund, N. E. (1998). *Assessment of student achievement*. Allyn & Bacon Publishing.
- Harlen, W. (2007). Assessment of learning. *Assessment of Learning*, 1-176.
- Harlen, W. (2012). *Assessment for learning: A practical guide*. Crown House Publishing.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385-402.
- Jalilzadeh, K., Dastgoshadeh, A., & Khosravi, R. (2023). The Impact of an In-Service Professional Development Course on Teachers Language Assessment Knowledge. *Journal of Modern Research in English Language Studies*. <https://doi.org/10.30479/jmrels.2023.18114.2152>
- Joughin, G. (2008). Assessment, learning and judgment in higher education: A critical review. *Assessment, learning and judgment in higher education*, 1-15.
- Kahl, S. R., Hofman, P., & Bryant, S. (2012). Assessment literacy standards and performance measures for teacher candidates and practicing teachers. *Assessment*.
- Leung, C. (2005). From formative assessment to assessment for learning: Exploring teachers' beliefs and practices. *Assessment in Education: Principles, Policy & Practice*, 12(2), 169-182.
- Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329-344.
- McMillan, J. H. (2000). Fundamental assessment principles for teachers and school administrators. *Practical Assessment, Research, and Evaluation*, 7(1), 8.
- McMillan, J. H. (2014). *Classroom assessment: Principles and practice for effective standards-based instruction*. Pearson.
- Naraghizadeh, M., Azizmalayeri, F., & Khalaji, H. (2022). EFL Teachers' Perceptions and Practices of Alternative Assessment Strategies and Their Relationship with Teacher Reflection. *Journal of English Language Pedagogy and Practice*, 15(30), 266-290. DOI: 10.30495/jal.2022.697012
- Nunan, D., & Bailey, K. M. (2009). *Exploring second language classroom research: A comprehensive guide*. Boston, Heinle CENGAGE Learning.
- O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing*, 30(3), 363-380.
- Pat-El, R., et al. (2013). The power of formative assessment to advance learning and teaching in primary schools: A case study from Israel. *The Curriculum Journal*, 24(1), 5-29.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). The nature of assessment and reasoning from evidence. *Knowing what students know: The science and design of educational assessment*, 37-54.
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381-402.
- Popham, W. J. (2006). Assessment for Learning: An Endangered Species? *Educational Leadership*, 63(5), 82-83.



- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 48(1), 4-11.
- Popham, W. J. (2017). *Assessment literacy for all teachers and administrators: It's not just about testing!* Routledge.
- Price, M., Rust, C., O. Donovan, B., Handley, K., & Bryant, R. (2012). *Assessment literacy: The foundation for improving student learning*. ASKe, Oxford Centre for Staff and Learning Development.
- Shohamy, E. (2001). Democratic assessment as an alternative. *Language testing*, 18(4), 373-391.
- Stiggins, R. (2008, September). Assessment for learning, the achievement gap, and truly effective schools. In *Educational Testing Service and College Board Conference, Portland, OR*.
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, 72(7), 534-539.
- Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238.
- Stiggins, R. J. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan*, 81(3), 191-198.
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758-765.
- Stiggins, R. J. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan*, 87(4), 324-328.
- Taylor, L. (2009). *Mourning Dress (Routledge Revivals): A Costume and Social History*. Routledge.
- Vogt, K., Sperling, I., & Brüstle, M. (2018). *TALE-eeeeer'' assessment literacy enhancement needs analysis report. Erasmus+ European Union. Project No. 2015-1CY01-KA201-011863 (Project No. 2015-1-CY01-KA201-011863)*.
- Vogt, K., Tsagari, D., Csépes, I., Green, A., & Sifakis, N. (2020). Linking learners' perspectives on language assessment practices to teachers' assessment literacy enhancement (TALE): Insights from four European countries. *Language Assessment Quarterly*, 17(4), 410-433.
- Walters, F. S. (2010). Cultivating assessment literacy: Standards evaluation through language-test specification reverse engineering. *Language Assessment Quarterly*, 7(4), 317-342.
- Wiggins, G., & McTighe, J. (1998). What is backward design? *Understanding by design*, 1, 7-19.
- William, D., and M. Thompson. 2008. Integrating assessment with learning: What will it take to make it work? In *The future of assessment: Shaping teaching and learning*, ed. C.A. Dwyer, 53-82. New York: Erlbaum.
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and teacher education*, 58, 149-162.