Research Article

Assessment Literacy and its Relevance to Teaching Experience and Academic Background among EFL Teachers in Iran

Rahim Watmani[1], *Hassan Assadollahfam[2], Bahram Behin[3]

[1]PhD Candidate of TEFL, English Language Department, Islamic Azad University, Bonab Branch, Iran, E-mail: rahimvatmani@gmail.com, https://orcid.org/0000-0001-5628-3896
[2]PhD in Applied Linguistics, Islamic Azad University, Bonab Branch, Iran, Email: h_asadollahfam@yahoo.com, https://orcid.org/: 0000-0003-1643-3372
[3]Associate Professor in ELT, Azarbaijan Shahid Madani University, Tabriz, Iran, Email: E-mail: bahram.behin@gmail.com, https://orcid.org/: 0000-0002-2224-815X
*Corresponding author: Hassan Assadollahfam, Email: h_asadollahfam@yahoo.com

ABSTRACT

Teachers' educational background and experience in classroom assessment are important factors that influence their level of assessment literacy. Despite this, few studies are reported to investigate this area. The present study aims to find differences between experienced and novice teachers with or without TEFL backgrounds regarding their assessment literacy skills. Two hundred high school teachers working in various Iranian schools with TEFL and non-TEFL backgrounds were randomly selected to participate in the current study. Chi-square (crosstabs) was applied to compare the assessment practices of the study groups. The researchers also took notes while observing the classroom assessments. The results revealed that TEFL-background teachers practiced assessment methods more significantly than the teachers with no TEFL background. Likewise, it was indicated that novice teachers practiced assessment methods and components more significantly than experienced teachers, irrespective of their educational background. The findings also indicated that high school teachers of TEFL have inadequate educational assessment literacy. Therefore, ongoing in-service training programs on educational assessment should be considered to cater to problems of low assessment literacy in Iranian high schools.
Keywords: assessment literacy, assessment literacy practices, novice and experienced teachers, TEFL/non-TEFL background

1. INTRODUCTION

Assessment is considered an integral part of second language teaching and learning because it can affect both the quality of instruction presented by the teacher and the quality of learning on the student's side. While this is the case, a collection of research findings shows that teachers struggle to integrate assessment with instruction in alignment with contemporary assessment theories and principles (Amiri & Birjandi, 2015; Bachman & Palmer, 2010; DeLuca & Klinger, 2010; Scarino, 2013; Lam, 2019). Specifically, previous research has shown that Iranian EFL teachers have an inadequate knowledge of assessment literacy (Ashraf & Zolfaghari, 2018; Fard & Tabatabaei, 2018; Farhady & Tavassoli, 2018; Mellati & Khademi, 2018; Tavassoli & Farhady, 2018). Besides, Razavipour *et al.* (2011) argue that educators need a better grasp of assessment and adjust their ways of teaching and assessing English to satisfy standardized exam demands.

According to the reformed assessment policies in the Iranian ELT context (Firoozi *et al.*, 2019), only teachers with a TEFL background can teach in high schools. The Ministry of Education regulations have confirmed the same notion (Farhady & Tavassoli, 2018). Nevertheless, many Iranian high schools prefer to avoid trained teachers in TEFL and are familiar with AL principles (Ashraf & Zolfaghari, 2018; Mellati & Khademi, 2018). As AL plays a significant role in the teachers' reflective teaching (Latif, 2021; Tajeddin *et al*., 2022; Tavassoli & Farhady, 2018) and practical training (Ellis, 2008), employing teachers with a TEFL background instead of the ones without a TEFL background is expected. Another significant point to consider is that the assessment practices of EFL teachers in the Iranian context need to receive more attention in the research agenda. Hence, the present study aims to demystify the assessment practices of experienced and novice teachers with and without a TEFL background in Iranian high schools.

2. REVIEW OF LITERATURE

2.1 Novice and Experienced Teachers

Novice teachers have at most 2 years of teaching experience (Gatbonton, 2008). Identifying experienced

teachers is much more complicated. Experienced teachers have many years of teaching experience, know how to manage their classrooms effectively, and can motivate students.

Rodríguez and McKay (2010, p. 2) present attributes of experienced ESL teachers that distinguish them from novice teachers. They present that such teachers have more control over classroom routines and are less concerned about students' adverse reactions to class activities or the learning process. They are English teachers with more than 3 years of experience.

In the present study, English teachers who have formally studied one of the branches of the English language, such as English literature, TEFL, or translation studies at home or abroad at the B. A or M.A. levels are considered as teachers with TEFL backgrounds. On the other hand, the teachers who have studied other majors and have developed enough English and can teach this language in Iranian high schools are labeled as teachers with non-TEFL backgrounds. In the present study, teachers with an English educational background and those without it were considered.

2.2 International Studies on Assessment Literacy

Some researchers have investigated the multiple aspects of teacher assessment literacy, considering how it works, what it implies, and how teacher training programs can promote Teacher Assessment Literacy (TAL) (e.g., Barnes, Fives, and Dacey, 2015; Braden *et al.*, 2005; Burry-Stock & Frazier, 2008; Farhady & Tavassoli, 2018; Khodabakhshzadeh et al., 2018; Mertler, 2005; Stiggins, 1991). In particular, studies on TAL have been accomplished in second-language literature since the 1990s (Stiggins, 1991). For example, a group of researchers has investigated the assessment literacy practices of EFL teachers in the classroom (Crusan et al., 2016; Lam, 2019; Sanders & Rivers, 1996; Vitali, 1994). Some studies have also examined EFL instructors' instructional strategies relevant to AL practices (Eslami & Fatahi, 2008; Hajizadeh & Salahshour, 2014; Razavipour et al., 2011). Others have found that TAL is a major teacher qualification in the EFL domain (Ashraf & Zolfaghari, 2018; Bayat & Rezaei, 2015; Zamani & Ahangari, 2016). Still, other researchers have concluded that deficiency in AL may create problems for EFL teachers in their lesson

planning (Stobart, 2008). Likewise, Ellis (2008) found that EFL teachers with less interest in the assessment are likely to be strict and create an atmosphere of competition for their learners, negatively affecting the student's learning and second language development. Inbar-Lourie, 2013 and Popham, 2014 claim that AL is the key to effective teaching. James (2008) found that AL can be crucial in evaluating language teaching mechanisms and assessment as learning, assessment for learning, and dynamic assessment. Taylor (2013) found that teachers' assessment awareness is significant for teachers, and as such, language educators should emphasize this in their teacher education programs (Mellati & Khademi, 2018; Tavassoli & Farhady, 2018; Taylor, 2013). A plethora of studies have been conducted on AL as an outcome of Professional Development (PD) or teacher training programs (Atay, 2008; Firoozi et al., 2019; Looney *et al.*, 2018; Walters, 2010; Wang et al., 2008), and some studies have also looked at assessment literacy as a notion related to students' achievement (Birjandi & Tamjid, 2012; Braney, 2010). AL has also been measured by comparing the AL of different teachers, such as novice and experienced teachers (Edwards, 2017; Gatbonton, 2008; Mertler, 2005; Tajeddin et al., 2018). Moreover, Azadi (2018) investigated the conceptual factors of TAL among Iranian English for Special Purposes (ESP) instructors and Mohammadi (2020) compared ESP instructors' AL knowledge with TEFL and non-TEFL backgrounds. Likewise, Jeong (2013) argues that the definition of AL differs for language and non-language testers. The overall findings of these studies show that teachers' AL practices do not necessarily represent their AL knowledge (Brookhart, 2011; Fard & Tabatabaei, 2018; Jannati, 2015; Lam, 2019; Mertler, 1999).

2.3 Studies on Assessment Literacy in Iran

Past research on AL has shown that Iranian EFL teachers' level of AL knowledge is inadequate (Ashraf & Zolfaghari, 2018; Fard & Tabatabaei, 2018; Farhady & Tavassoli, 2018; Mellati & Khademi, 2018; Tavassoli & Farhady, 2018). Moreover, EFL teachers' perceptions of Language Assessment Literacy (LAL) are not presented in their practices (Jannati, 2015). Babai Shishavan and Sadeghi (2019) found that AL can be considered one of the significant features of EFL teachers. Bayat and Rezaei (2015) also concluded that

one of the most crucial duties of instructors is student assessment since the caliber of instruction in the classroom is directly related to the caliber of the applied assessment. Jannati (2015) found that their practices do not reflect ELT teachers' language assessment literacy perceptions. Similarly, Mellati and Khademi (2018) insist that language educators consider teachers' assessment awareness in their teacher education programs. Tavassoli and Farhady (2018) found that recent advancements in education urge teachers to be aware of and implement effective methods of instruction and evaluation to improve students' learning. To tackle the current challenges, teacher education programs must provide opportunities for instructors to increase their expertise in several areas, including L2 assessment literacy.

Since many teachers teaching English at high schools do not enjoy a TEFL background (Farhady & Hedayati, 2009), the study attempts to consider this point and hence focuses on investigating the AL knowledge of English teachers with both TEFL and non-TEFL backgrounds. One of the areas left untouched in Iranian ELT research is the assessment literacy of EFL teachers working in high schools. Nor has their practical knowledge in this domain been taken into consideration by the ELT researchers.

Likewise, as experience plays a significant role in the development of AL (Eezami, 2016; Gatbonton, 2008; Mertler, 2003), the present study focuses on the difference between experienced and novice EFL teachers' AL both from the perspective of their theoretical knowledge and what they do in practice. Hence, the following questions are posed in the present study.

1. To what extent do assessment practices of EFL teachers with a TEFL background differ from those with a non-TEFL background?

2. To what extent do assessment practices of experienced and novice EFL teachers differ?

3. METHODOLOGY

3.1. Design

The present study employed a *mixed-methods sequential explanatory design* based on Creswell & Clark (2017). According to the principles of sequential explanatory design, the quantitative data were collected

and analyzed first, followed by the qualitative data in the second phase (Creswell, 2009). Furthermore, the qualitative data supported the findings from the first phase of the study and illustrated the unexplored parts of the study.

3.2. Participants

The participants of this study were 200 novices (n=120) and experienced (n=80) English teachers with TEFL (n=88) and non-TEFL (n=112) backgrounds, including both genders (males=85, females=115), with university education (B.A.,(n=175) or M.A.(n=25) degrees) teaching in different Iranian high schools in East and West Azerbaijan provinces. The participants' ages ranged from 25 to 50 or more. The selection process was based on teachers' willingness to participate in the study. They were selected based on stratified random sampling, "a method of sampling in which the researcher divides the population into subgroups known as strata and randomly selects participants from each group" (Creswell & Clark, 2017, p. 235).

Eighty-eight teachers taking part in the study were selected out of those with TEFL backgrounds and 112 from those with non-TEFL backgrounds teaching English at Iranian high schools.

Also, in both groups of teachers (TEFL and Non-TEFL) teaching English, two groups of novice and experienced teachers were focused on. Moreover, based on purposive sampling, 20 teacher participants were picked for a classroom observation. According to Mackey and Gass (2016), purposive sampling is a nonrandom type of sampling through which the researcher singles out some participants based on a set of criteria. The criteria for selecting the interviewees were their questionnaire scores, educational background, and teaching experience, provided that those selected teachers had agreed to further cooperation; otherwise, other teachers were selected instead. Table 1 shows the demographic information concerning the participants of the study. As indicated in the table, 85 participants were male, and 115 were female.

Table 1
Distribution of Participants by Gender

		Frequency	Percent
Gender	Male	85	42.50
	Female	115	57.50

	Total	200	100.0

In terms of the participants' background, Table 2 shows that 88 participants enjoyed the TEFL background, while 112 participants had a non-TEFL background.

Table 2

Distribution of Participants by Background

		Frequency	Percent
	TEFL	88	44
Background	Non-TEFL	112	56
	Total	200	100.0

In terms of the participants' teaching experience, Table 3 shows 120 participants were novice (less experienced), while 80 were experienced.

Table 3

Distribution of Participants by Teaching Experience

		Frequency	Percent
	Novice	120	60
	Experienced	80	40
Teaching Experience			100.0
	Total	200	

3.3. Instruments and Materials

3.3.1 Novice and Experienced Teacher Questionnaire

Rodríguez and McKay (2010) initially developed and validated this questionnaire, and its modified version considers the cultural and local notions already used in the Iranian context (Baniali, 2018; Eezami, 2016). The novice and Experienced Teacher Questionnaire can relatively specify the teachers' experience level. It consists of 12 items and five choices based on the Likert scale. The reliability index (Cronbnach's alpha) of the original version of the questionnaire is "(α=0.72). In the modified version, the reliability indices for this scale have been reported as α=0.76 (Eezami, 2016) and α=0.71(Baniali, 2018). The minimum and

maximum scores are 12 and 60, respectively, indicating that teachers below 30 were associated with novice teachers, and the ones with scores above 36 are considered experienced teachers.

3.3.2. Classroom Observation

To collect data on EFL teachers' practices of AL, the researcher used *observational note-taking*, which represented observers' inferences during the data collection process. Relying on observational data is usually suggested in qualitative studies as it entails observing and documenting genuine participant activities and familiarizing the researcher with the participant's classroom behaviors (FitzGerald & Mills, 2022).

The notes were evaluated based on the seven components of AL such as 1) appropriate assessment method selection, 2) developing proper assessment methods, 3) interpreting the assessment results, 4) making decisions in line with assessment results, 5) developing a reliable and valid scoring method, 6) discussing assessment results with learners' parents, and seven being conscious about ethical concepts of L2 assessment. Each standard was measured through 4 items in the checklist in the appendix section of this paper.

To estimate the reliability index of the observational notes, the researchers compared the observation notes and recorded sessions.

After obtaining the consent of the participants and authorities, the researcher observed and took notes of twenty classes, each lasting three sessions. Five classes belonged to experienced TEFL teachers, and five others to non-experienced non-TEFL teachers.

Each session took about 90 minutes. Indeed, within the present study's focus, observations allowed the researcher to analyze teachers' assessments in natural settings.

Table 4.

Intra-class Correlation Coefficients (Main Study)

		Intra-class Correlation	95% Confidence Interval		F Test with True Value			
			Lower Bound	Upper Bound	Value	df1	df2	Sig
Choosing Appropriate Assessment Methods	Single Measures	.589[a]	.192	.933	8.174	4	16	.001
	Average Measures	.878[c]	.544	.986	8.174	4	16	.001

Developing Appropriate	Single Measures	.407[a]	-.023	.888	3.750	4	12 .033
Assessment Methods	Average Measures	.733[c]	-.099	.970	3.750	4	12 .033
Administering, Scoring,	Single Measures	.818[a]	.515	.976	23.500	4	16 .000
Interpreting Results	Average Measures	.957[c]	.841	.995	23.500	4	16 .000
Using Assessment Results to Make	Single Measures	.800[a]	.481	.973	21.000	4	16 .000
Decisions	Average Measures	.952[c]	.822	.994	21.000	4	16 .000
Developing a Valid Grading	Single Measures	.450[a]	.068	.896	5.091	4	16 .008
Procedure	Average Measures	.804[c]	.267	.977	5.091	4	16 .008
Communicating Assessment	Single Measures	.435[a]	.063	.769	5.035	4	16 .006
Results	Average Measures	.705[c]	.065	.767	5.035	4	16 .006
Recognizing Unethical or Illegal	Single Measures	.385[a]	.021	.874	4.125	4	16 .017
Practices	Average Measures	.758[c]	.096	.972	4.125	4	16 .017

Note. The first and second rows are intra-rater and inter-rater reliability estimates.

3.4. Data Collection and Analysis Procedures

Several steps were taken to carry out the present study. In the first step, the 200 EFL teachers were selected based on stratified random sampling from both novices (n=120). They experienced (n=80) English teachers with TEFL and non-TEFL backgrounds, including both genders (males=85, females=115), with university education (B.A., (n=175) or M.A. (n=25) degrees) teaching in different Iranian high schools in West Azerbaijan province.

In the second step, the study's data collection instruments were prepared: the Teacher Assessment Literacy (Mertler, 2002) and the Novice and Experienced Teacher Questionnaire (Rodríguez & McKay, 2010). These instruments had already been used in the Iranian scholars' studies, and their reliability and validity had been confirmed.

In the third step, the two sets of questionnaires (Teacher Assessment Literacy and Novice and Experienced Teacher Questionnaire) were distributed among 200 English teachers with TEFL or non-TEFL

backgrounds in various high schools where English classes were being held. The teachers completed the survey during non-instructional times at their convenience, and their responses and identities were guaranteed to be kept confidential. Then, questionnaires were scored, and statistical methods were used to analyze the quantitative data.

In the qualitative phase of the study, after administering the two questionnaires and conducting data analysis, the researchers selected 20 teachers for classroom observations. The teachers were observed three times. Then, the average of the scores given to the teachers' behavior concerning their practices in terms of assessment in the three sessions observed for each class was taken into consideration. Also, the researcher made some inferences during the data collection process and took notes.

A Multivariate Analysis of Variance (MANOVA) was run in line with related research questions to determine the difference between novice and experienced English teachers with TEFL and non-TEFL backgrounds. Likewise, chi-square (crosstabs) was applied to compare the TEFL and non-TEFL teachers' assessment practices. The same was done to compare novice and experienced EFL teachers.

In the second phase, a content analysis was conducted to answer the study's qualitative research questions. It should be noted that all notes were categorized and later analyzed, and this categorization process led to the identification of the predominant patterns presented by the teachers. Later, themes were coded and analyzed.

4. RESULTS

The results were compared to the data collected through the questionnaires and the notes taken by the observers in the classroom sessions. In line with FitzGerald and Mills (2022), this was done to help standardize and streamline the data to increase the study's internal validity. As such, the quantitative results of the study are presented in line with the two research questions, followed by a description of the qualitative findings.

The first research question aimed to find the likely differences between assessment practices of English

teachers with TEFL and non-TEFL backgrounds, regardless of their experience. For the quantitative part of the study, chi-square (crosstabs) was applied to analyze the possible differences in the TEFL and non-TEFL teachers' assessment practices. The percentages, frequencies, and standardized residuals (Std. Residuals) for the TEFL and non-TEFL teachers' practice of assessment methods are displayed in Table 5. A residual higher than +/- 1.96, shows that the relevant option was selected more/less significantly than what was expected. The results showed that TEFL teachers selected "very much" (Std. Residual = 3 > 1.96) more than the non-TEFL teachers, while non-TEFL teachers selected "little" (Std. Residual = 4.4 > 1.96) and "a little" (Std. Residual = 4.2 > 1.96) more than teachers with TEFL background. In line with these results, it can be concluded that TEFL practiced assessment methods more significantly than non-TEFL teachers.

Table 5
Frequencies, Percentages, and Std. Residuals; Assessment Practice by Background

		Little	A little	To some extent	Much	Very much	Total
				Choices			Total
TEFL	Count	0	1	58	124	97	280
	%	0.0%	0.4%	20.7%	44.3%	34.6%	100.0%
	Std. Residual	-4.4	-4.2	-.4	1.5	3.0	
Non-TEFL	Count	38	38	65	93	46	280
	%	13.6%	13.6%	23.2%	33.2%	16.4%	100.0%
	Std. Residual	4.4	4.2	.4	-1.5	-3.0	
Total	Count	38	39	123	217	143	560
	%	6.8%	7.0%	22.0%	38.8%	25.5%	100.0%
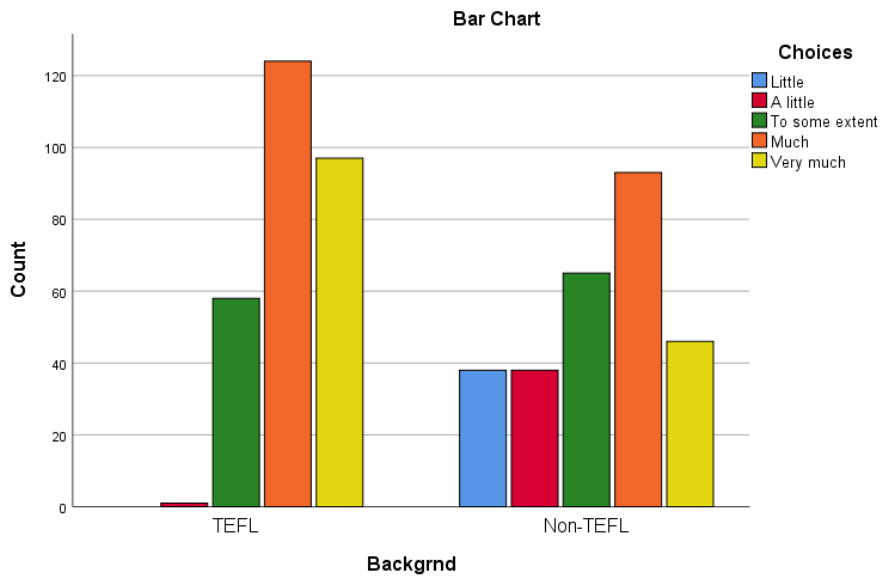
Figure 1 *Frequencies of Assessment Practice by Background*

The results of the chi-square analysis (χ^2 (4) = 96.11, p = .000, Cramer's V= .414, displaying a moderate to large effect size) (Table 6) showed that TEFL teachers practiced assessment methods more significantly than non-TEFL teachers.

Table 6
Chi-Square Tests: Assessment Practice by Background

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	96.118ᵃ	4	.000
Likelihood Ratio	120.881	4	.000
Linear-by-Linear Association	84.941	1	.000
N of Valid Cases	560		
Cramer's V	.414		.000

b. 0 cells (00.0%) have an expected count of less than 5. The minimum expected count is 19.

The second research question was an attempt to find the differences between the assessment practices of experienced and novice English teachers, irrespective of their educational background. For this, the researchers relied on the frequencies, percentages, and std. Residuals of assessment practices. The frequencies, percentages, and standardized residuals (Std. Residuals) for the novice and experienced teachers' practice of assessment methods are displayed in Table 7. The results indicated that novice teachers

selected "very much" (Std. Residual = 3 > 1.96) more than the experienced teachers, while experienced teachers selected "little" (Std. Residual = 4.4 > 1.96) and "a little" (Std. Residual = 4.2 > 1.96) more than novice teachers. In line with these results, novice teachers practiced assessment methods and components more significantly than experienced teachers, regardless of their educational background.

The results of the analysis of Chi-square (χ^2 (4) = 92.04, p = .000, Cramer's V = .405 representing a moderate to large effect size) (Table 8) indicated that the novice teachers practiced assessment methods and components more significantly than experienced teachers.

Table 7
Frequencies, Percentages, and Std. Residuals; Assessment Practice by Teaching Experience

		Choices					Total
		Little	A little	To some extent	Much	Very much	
Experienced	Count	38	38	43	88	73	280
	%	13.6%	13.6%	15.4%	31.4%	26.1%	100.0%
	Std. Residual	4.4	4.2	-2.4	-2.0	.2	
Novice	Count	0	1	80	129	70	280
	%	0.0%	0.4%	28.6%	46.1%	25.0%	100.0%
	Std. Residual	-4.4	-4.2	2.4	2.0	-.2	
Total	Count	38	39	123	217	143	560
	%	6.8%	7.0%	22.0%	38.8%	25.5%	100.0%
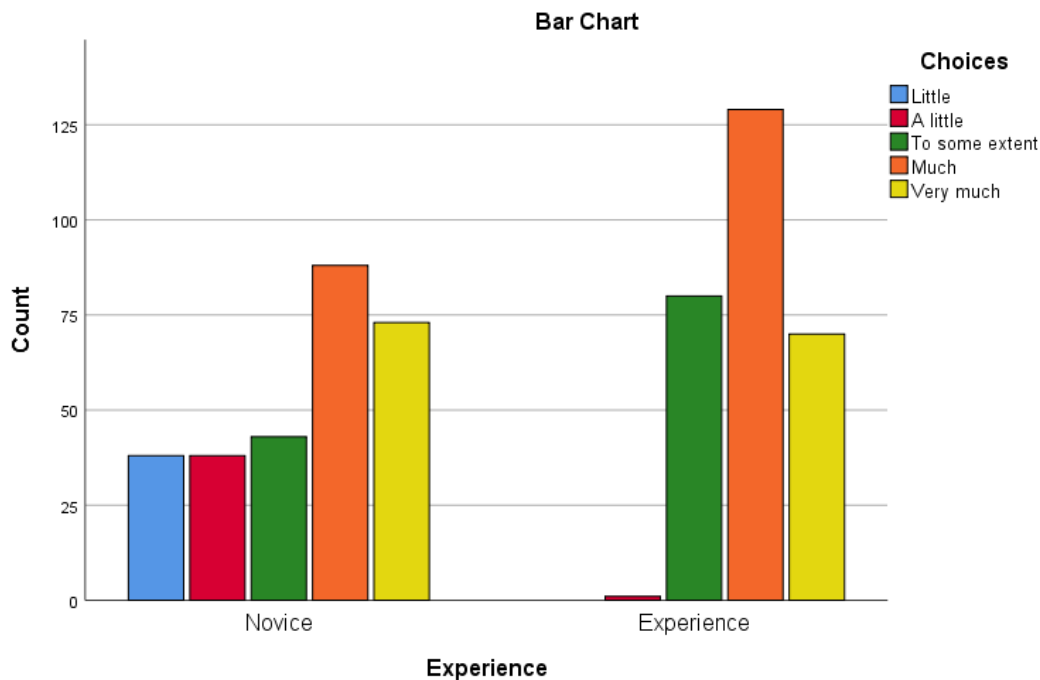
Figure 2 *Frequencies of Assessment Practice by Teaching Experience*

Table 8

Chi-Square Tests; Assessment Practice by Teaching Experience

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	92.042[a]	4	.000
Likelihood Ratio	116.604	4	.000
Linear-by-Linear Association	30.745	1	.000
N of Valid Cases	560		
Cramer's V	.405		.000

b. 0 cells (00.0%) have an expected count of less than 5. The minimum expected count is 19.

   In the qualitative phase of the study, twenty teachers were observed three times. The researcher wanted to see if there are any differences between novice and experienced teachers with or without TEFL backgrounds. For this purpose, he took some notes based on his inferences. The main points of the notes the researcher had taken and his inferences concerning the differences between assessment practices of experienced and novice English teachers with and without a TEFL background were as follows:

1. Both novice and experienced TEFL teachers could choose more appropriate assessment methods than non-TEFL teachers.

2. Novice (TEFL) teachers were the most knowledgeable in developing, designing and realizing the Significance of standardized tests.

3. TEFL teachers were well-managed and could handle the job more accurately than their non-TEFL counterparts regarding test administration factors.

4. Experienced teachers with a TEFL background were more successful than their non-TEFL counterparts in using assessment results to make decisions,. Novice TEFL teachers paid special attention to the prerequisite knowledge of their students before teaching the new unit. They were also the only group that was relatively aware of concepts such as norm-referenced and criterion-referenced and could use the information they had to plan for their instructions.

5. The experienced teachers of both groups and the novice teachers with a TEFL background behaved appropriately regarding grading methods, communicating the assessment results and ethical aspects of L2 assessment. Overall, novice teachers with a TEFL background practiced assessment methods and components more significantly than experienced teachers with or without a TEFL background.

5. DISCUSSION

The results of data analysis firstly indicated that the teachers who had studied TEFL at university practiced assessment methods more significantly than the ones without a TEFL background. This supports the results of many studies regarding AL practices in the L2 classroom (Eslami & Fatahi, 2008; Hajizadeh & Salahshour, 2014; Razavipour *et al.*, 2011). The studies above mainly consider AL practices as EFL instructors' instructional strategies and argue that with AL literacy, the EFL teacher can be as effective as expected in educational norms (Ellis, 2008). Likewise, the present finding, which signifies the priority of the EFL teachers with a TEFL background over the non-TEFL teachers teaching English in high schools, can take support from Hajizadeh and Salahshour's (2014) study confirming that AL is an indispensable feature of effective EFL instructors. In this respect, the present finding can also take support from some international studies done on the EFL teachers' AL practices (Black *et al.*, 2003; Braden *et al.*, 2005; Brookhart, 2011; Burry-Stock & Frazier, 2008; Crusan *et al.*, 2016; James, 2008; Lam, 2019; Looney *et al.*, 2018; McTighe & O'Connor, 2005; Mertler, 1999; Pilcher, 2001; Popham, 2014; Xu & Brown, 2016) as well as a couple of studies carried out on AL practices and TEFL teachers in the Iranian context (Ashraf & Zolfaghari, 2018; Birjandi & Tamjid, 2012; Eslami & Fatahi, 2008; Fard & Tabatabaei, 2018; Farhady & Tavassoli, 2018; Firoozi *et al.*, 2019; Hajizadeh & Salahshour, 2014; Khodabakhshzadeh *et al.*, 2018; Razavipour *et al.*, 2011; Tavassoli & Farhady, 2018). Moreover, the present finding aligns with Alemi, Miri, and Mozafarnezhad (2019) and Azadi's (2018) study exploring the conceptual factors of TAL among Iranian ESP instructors. Likewise, the present study finding is in concordance with Jeong (2013), who

argues that the definition of AL differs for assessors with a language education background and those with a non-language education, as language teachers with a TEFL background are more familiar with the notions and constructs to be tested and assessed.

Secondly, the results revealed that novice teachers practiced assessment methods and components more significantly than experienced teachers. This is not in line with most of the previous studies comparing the AL of novice and experienced teachers (Braden *et al.*, 2005; Edwards, 2017; Gatbonton, 2008; Mertler, 2005) as the previous studies mainly mention that experienced teachers enjoy more relevant AL practices in the L2 classroom. The reason might lie in the sensitivity to teaching assessment notions and principles to the TEFL students in the teacher training programs in the Iranian context on the one hand (Farhady & Tavassoli, 2018; Mohammadi, 2020), and teacher burnout issues among the experienced ones (Fathi & Derakhshan, 2019; Fathi & Saeedian, 2020; Khezerlou, 2017; Roohani & Dayeri, 2019; Sadeghi & Khezrlou, 2014), on the other hand. In this respect, Tajeddin *et al.* (2018) found that irrespective of the moderate divergence that exists between the beliefs of less experienced and more experienced teachers about AL standards, both need to update their assessment strategies and practices.

However, in terms of the success of novice EFL teachers in applying AL principles in their classroom practices, the preset study's findings are similar to Mertler's (1999) findings regarding Ohio teachers' assessment literacy in practice. The present findings also confirm the findings of Pilcher (2001), who presented AL standards, and those of Xu and Brown (2016), who studied teacher assessment literacy in practice in the Chinese context. Like the present study findings, these studies confirm that new university graduates are more interested in employing their achievements in the L2 classroom when compared to experienced teachers.

6. CONCLUSION

The present study's findings showed that novice teachers who recently graduated from universities are more knowledgeable than their experienced counterparts regarding AL practices (Mohammadi, 2020). They are

also more sensitive to notions of assessment literacy (Farhady & Tavassoli, 2018). The reason might lie in the teacher burnout issues with the experienced teachers (Fathi & Derakhshan, 2019; Sadeghi & Khezrlou, 2014) and the interest of new TEFL graduates in using their achievement in the L2 classroom. After all, results revealed that all EFL teachers involved in the study, especially novice teachers without a TEFL background and experienced ones, urgently needed in-service training courses in L2 assessment concepts. In general, it was found that the Iranian EFL teachers involved in the study, regardless of their experience and background, needed to be more practically familiar with the standards of teacher competencies presented by Plake and Impara (1993). Nor were they well familiar with the standards of assessment literacy (Mertler, 2003) or AL practices presented by other researchers (McTighe & O'Connor, 2005; Pilcher, 2001), and this gets more troublesome when considering that almost half of these teachers who have a background of teaching EFL at the high school level enjoy a TEFL background. Indeed, they have passed courses in language testing and assessment at their undergraduate and graduate levels. This aligns with Lam's (2015) suggestion on language assessment training for ESL teachers in Hong Kong and its effect on the L2 learners' development. Hence, an in-service training course in assessment literacy is suggested for the teachers taking part in the study and many others active in Iranian high schools. Specifically, they must understand that assessment is an integral component of instruction and goals for student learning (McMillan, 2000; Pilcher, 2001). Teachers have also mentioned that they need to get acquainted with the application of assessment literacy concepts more than the fundamental considerations of measurement principles (Rogers, 1991). As such, teacher educators involved in teaching assessment and measurement courses can teach EFL teachers the principles of AL and illuminate this vital connection between assessment and instruction, making assessment more applicable to teaching principles.

This study's findings have some implications too. The findings imply that EFL teachers' knowledge of assessment needs to significantly improve in communicating assessment results to others and validating the learners' assessment. Therefore, to address the problems of teachers with low levels of AL, ongoing in-

service training programs on educational assessment and assessment literacy are urgent. Although TEFL high school teachers had taken assessment and language testing courses in this study, more than one course in assessment and measurement is required to cover everything that high school teachers need to know.

The findings of the present study may offer some suggestions for further research. Future studies might consider analyzing the impact of EFL teachers' in-service and pre-service instruction in terms of AL. A semi-longitudinal study of the concept of AL on university professors with and without TEFL backgrounds can illuminate the effect of the longitudinal effect of instruction on AL. Finally, further research is recommended to explore the role of AL instruction in developing teachers' motivation, teachers' self-regulatory factors of teaching and learning, and their relationship with their autonomy.

REFERENCES

Alemi, M., Miri, M., & Mozafarnezhad, A. (2019). Investigating the effects of online concurrent group dynamic assessment on enhancing grammatical accuracy of EFL learners. *International Journal of Language Testing, 9*(2), 29-43.

Amiri, M., & Birjandi, P. (2015). Reliability and content validity of a comprehensive discourse completion test of interlanguage pragmatics for academic situations. *Sino-US English Teaching, 12*(9), 654-659. http://dx.doi.org/10.17265/1539-8072/2015.09.002

Arani, A. M., Kakia, M. L., & Karimi, M. V. (2012). Assessment in education in Iran. *Assessment*, 9(2), 101-110.

Ashraf, H., & Zolfaghari, S. (2018). EFL teachers' assessment literacy and their reflective teaching. *International Journal of Instruction, 11*(1), 425-436. https://doi.org/10.12973/iji.2018.11129a

Atay, D. (2008). Teacher research for professional development. *ELT Journal, 62*(2), 139-147. https://doi.org/10.1093/elt/ccm041

Azadi, A. (2018). *A study on the conceptual factors of teacher assessment literacy among ESP instructors* (Unpublished master's thesis). Islamic Azad University, Electronic Branch, Tehran, Iran.

Babai Shishavan, H., & Sadeghi, K. (2009). Characteristics of effective English language teacher as perceived by Iranian teachers and learners of English. *Iranian Journal of Language Teaching Research, 1*(2), 130-132.

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice*. OUP.

Baniali, S. (2018). *A study on Iranian experienced and novice EFL teachers' belief and practice in teaching vocabulary* (Unpublished master's thesis). North Tehran Branch, Islamic Azad University, Tehran, Iran.

Barnes, N., Fives, H., & Dacey, C. M. (2015). Teachers' beliefs about assessment. In H. Fives & G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 284-300). Routledge.

Bayat, K., & Rezaei, A. (2015). Importance of teachers' assessment literacy. *International Journal of English Language Education, 3*(1), 139-146. https://doi.org/10.5296/ijele.v3i1.7260

Birjandi, P., & Tamjid, N. H. (2012). The role of self-, peer, and teacher assessment in promoting Iranian EFL learners' writing performance. *Assessment and Evaluation in Higher Education, 37*(5), 513-533. https://doi.org/10.1080/02602938.2010.549204

Braden, J. P., Huai, N., White, J. L., & Elliot, S. N. (2005). Effective professional development to support inclusive large-scale assessment practices for all children. *Assessment for Effective Intervention, 30*(4), 63-71. https://doi.org/10.1177/073724770503000407

Barney, B. T. (2010). *An examination of fourth-grade teachers' assessment literacy and its relationship to students' reading achievement* [Master's thesis, East Carolina University]. ProQuest Dissertations and Theses Global.

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practices, 30*(1), 3-12. https://doi.org/10.1111/j.1745-3992.2010.00195.x

Burry-Stock, J. A., & Frazier, C. H. (2008, March). *Revision of the assessment practice inventory (APIR): A combined exploratory factor analysis and polytomous IRT approach* [Paper presentation]. American Educational Research Association Annual Meeting, New York, NY, United States.

Chen, C. I. (2013). *Examining psychometric dimensions of the Ages and Stages Questionnaires: Inventory: A cross-country comparison between Taiwan and the United States* [Doctoral dissertation, University of Oregon]. Scholars' Bank. https://scholarsbank.uoregon.edu/xmlui/handle/1794/13115

Creswell, J. W., & Hirose, M. (2019). Mixed methods and survey research in family medicine and community health. *Family Medicine and Community Health, 7*(2), e000086. https://doi.org/10.1136/fmch-2018-000086

Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing, 28*, 43-56. https://doi.org/10.1016/j.asw.2016.03.001

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1). https://doi.org/10.14507/epaa.v8n1.2000

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

DeLuca, C., & Volante, L. (2016). Assessment for learning in teacher education programs: Navigating the juxtaposition of theory and praxis. *Journal of the International Society for Teacher Education, 20*(1), 19-31.

Eckhout, T., Davis, S., Mickelson, K., & Goodburn, A. (2005). A method for providing assessment training to in-service and pre-service teachers. In D. Gutiérrez (Ed.), *Proceedings of the annual meeting of the southwestern educational research association* (pp. 121-132). Arnolds.

Edwards, F. (2017). A rubric to track the development of secondary pre-service and novice teachers' summative assessment literacy. *Assessment in Education: Principles, Policy and Practice, 24*(2), 205-227. https://doi.org/10.1080/0969594X.2016.1225668

Eezami, R. (2016). *A study on work engagement and fulfillment of basic psychological needs among novice and experienced EFL teachers in the Iranian institutes* [Master's thesis, Kharazmi University]. http://thesis.khu.ac.ir/thesis/Thesis/ThesisView.aspx?ID=100

Ellis, R. (2008). *The study of second language acquisition (2nd ed.).* Oxford University Press.

Eslami, Z. R., & Fatahi, A. (2008). Teachers' sense of self-efficacy, English proficiency, and instructional strategies: A study of nonnative EFL teachers in Iran. *TESL-EJ, 11*(4), 1-19. http://tesl-ej.org/ej44/a1.html

Falsgraf, C. (2005). Why a national assessment summit? New visions in action. *National Assessment Summit, 3*, 6-9.

Fard, Z. R., & Tabatabaei, O. (2018). Investigating assessment literacy of EFL teachers in Iran. *Journal of Applied Linguistics and Language Research, 5*(3), 91-100. http://www.jallr.com/index.php/JALLR/article/view/780

Farhady, H., & Hedayati, H. (2009). Language assessment policy in Iran. *Annual Review of Applied Linguistics, 29*, 132-141. https://doi.org/10.1017/S0267190509090100

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), 79-94.

Fathi, J., & Derakhshan, A. (2019). Teacher self-efficacy and emotional regulation as predictors of teaching stress: An investigation of Iranian English language teachers. Teaching English Language, 13(2), 117-143. https://doi.org/10.22132/tel.2019.100527

Fathi, J., & Saeedian, A. (2020). A structural model of teacher self-efficacy, resilience, and burnout among

Iranian EFL teachers. *Iranian Journal of English for Academic Purposes, 9*(2), 14-28. https://doi.org/10.22034/ijceap.2020.240130.1144

Firoozi, T., Razavipour, K., & Ahmadi, A. (2019). The language assessment literacy needs of Iranian EFL teachers with a focus on reformed assessment policies. *Language Testing in Asia, 9*(1), 2-14. https://doi.org/10.1186/s40468-019-0080-5

FitzGerald, J., & Mills, J. (2022). The importance of ethnographic observation in grounded theory research. *Forum: Qualitative Social Research, 22*(2), 2-13. http://dx.doi.org/10.17169/fqs-22.2.3810

Gatbonton, E. (2008). Looking beyond teachers' classroom behavior: Novice and experienced ESL teachers' pedagogical knowledge. *Language Teaching Research, 12*(2), 161-182. https://doi.org/10.1177/1362168807086286

Hajizadeh, N., & Salahshour, N. (2014). Characteristics of effective EFL instructors: Language teachers' perceptions versus learners' perceptions. *International Journal of Applied Linguistics and English Literature, 3*(1), 202-214. https://doi.org/10.7575/aiac.ijalel.v.3n.1p.202

Inbar-Lourie, O. (2013). Language assessment literacy: What are the ingredients? Plenary speech at the 4th CBLA SIG Symposium Program' Language Assessment Literacy- LAL, Cyprus. Published in the special issue on language assessment literacy. *Language Testing, 30*(3), 301-307. https://doi.org/10.1177/0265532213480128

James, M. (2008). Assessment and learning. In G. Havnes & M. McDowell (Eds.), *Unlocking assessment: Understanding for reflection and application* (pp. 20-35). Routledge.

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

Jan-nesar, M. Q., Khodabakhshzadeh, H., & Motallebzadeh, K. (2020). Assessment literacy of Iranian EFL teachers: A review of recent studies. *Journal of Asia TEFL, 17*(2), 689-697. https://doi.org/10.18823/asiatefl.2020.17.2.17.689

Jeong, H. (2013). Defining assessment literacy: Is it different for language testers and non-language testers? *Language Testing, 30*(3), 345-362. https://doi.org/10.1177/0265532213480334

Khezerlou, E. (2017). Professional self-esteem as a predictor of teacher burnout across Iranian and Turkish EFL teachers. *Iranian Journal of Language Teaching Research, 5*(1), 113-130.

Khodabakhshzadeh, H., Kafi, Z., & Hosseinnia, M. (2018). Investigating EFL teachers' conceptions and literacy of formative assessment: Constructing and validating an inventory. *International Journal of Instruction, 11*(1), 139-152. https://doi.org/10.12973/iji.2018.11110a

Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment

literacy. *Language Testing, 32*(2), 169-197. https://doi.org/10.1177/0265532214547894

Lam, R. (2019). Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong. *System, 81*, 78-89. https://doi.org/10.1016/j.system.2019.01.005

Lantolf, J. P., & Poehner, M. E. (2008). Dynamic assessment. In E. Shohamy & N. Hornberger (Eds.), *Encyclopedia of language and education* (Vol. 7): Language testing and assessment (pp. 273-284). Springer.

Latif, M. W. (2021). Exploring tertiary EFL practitioners' knowledge base component of assessment literacy: Implications for teacher professional development. *Language Testing in Asia, 11*(1), 1-22. https://doi.org/10.1186/s40468-021-00123-8

Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2018). Re-conceptualizing the role of teachers as assessors: Teacher assessment identity. *Assessment in Education: Principles, Policy and Practice, 25*(5), 442-467. https://doi.org/10.1080/0969594X.2016.1268090

Mackey, A., & Gass, S. M. (2016). *Second language research: Methodology and design (2nd ed.).* Lawrence Erlbaum Associates.

McMillan, J. H. (2000). Fundamental assessment principles for teachers and school administrators. *Practical Assessment, Research and Evaluation, 7*(8), 1-9. https://scholarworks.umass.edu/pare/vol7/iss1/8

McNamara, T. (2003). Tearing us apart again: The paradigm wars and the search for validity. *Eurosla Yearbook, 3*(1), 229-238. https://doi.org/10.1075/eurosla.3.13mcn

McTighe, J., & O'Connor, K. (2005). Seven practices for effective teaching. *Educational Leadership, 63*(3), 10-17.

Mellati, M., & Khademi, M. (2018). Exploring teachers' assessment literacy: Impact on learners' writing achievements and implications for teacher development. *Australian Journal of Teacher Education, 43*(6), 1-18. https://doi.org/10.14221/ajte.2018v43n6.1

Mertler, C. A. (1999). Assessing student performance: A descriptive study of the classroom assessment practices of Ohio teachers. *Education, 120*, 285-296.

Mertler, C. A. (2002). *Classroom assessment literacy inventory. Adapted from the teacher assessment literacy questionnaire* (1993), by B. S. Plake & J. C. Impara, University of Nebraska-Lincoln, in cooperation with The National Council on Measurement in Education and the W.K. Kellogg Foundation.

Mertler, C. A. (2005). Secondary teachers' assessment literacy: Does classroom experience make a

difference? *American Secondary Education, 33*(2), 76-92.

Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools, 12*(1), 101-113. https://doi.org/10.1177/1365480208099288

Mohammadi, A. (2020). *A mixed-methods study on the teacher assessment literacy of ELT instructors versus content instructors Islamic Azad University* (Unpublished master's thesis). University of Qom, Iran.

Plake, B. S., & Impara, J. C. (1993). *Teacher assessment literacy questionnaire*. Nebraska-Lincoln: The National Council on Measurement in Education and the W.K. Kellogg Foundation.

Pilcher, J. K. (2001). *The standards and integrating instructional and assessment practices* (ERIC Document Reproduction Service No. ED 451 190).

Popham, W. J. (2014). *Classroom assessment: What teachers need to know (7th ed.).* Pearson Education.

Rezai, A., Alibakhshi, G., Farokhipour, S., & Miri, M. (2021). A phenomenographic study on language assessment literacy: Hearing from Iranian university teachers. *Language Testing in Asia, 11*(1), 1-25. https://doi.org/10.1186/s40468-021-00124-7

Razavipour, K., Riazi, A., & Rashidi, N. (2011). On the interaction of test washback and teacher assessment literacy: The case of Iranian EFL secondary school teachers. *English Language Teaching, 4*(1), 156-161. https://doi.org/10.5539/elt.v4n1p156

Rodríguez, A. G., & McKay, S. (2010). Professional development for experienced teachers working with adult English language learners. *CAELA Network Brief*. Retrieved from [4](https://www.cal.org/caelanetwork/resources/professionaldevelopment.html)

Roohani, A., & Dayeri, K. (2019). On the relationship between Iranian EFL teachers' burnout and motivation: A mixed methods study. *Iranian Journal of Language Teaching Research, 7*(1), 77-99.

Sadeghi, K., & Khezrlou, S. (2014). Burnout among English language teachers in Iran: Do socio-demographic characteristics matter? *Procedia-Social and Behavioral Sciences, 98*, 1590-1598. https://doi.org/10.1016/j.sbspro.2014.03.577

Sanders, W., & Rivers, J. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. http://mdk12.org/practices/ensure/tva/tva_2.html

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

Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Pre-service teachers' assessment literacy.

Journal of Science Teacher Education, 22(4), 371-391. https://doi.org/10.1007/s10972-011-9231-6

Stiggins, R. (1991). Assessment literacy. *Phi Delta Kappan, 3*(2), 534-539.

Stiggins, R. (1999). Are you assessment literate? *The High School Journal, 6*(5), 20-23.

Stobart, G. (2008). *Testing times: The uses and abuses of assessment.* Routledge.

Sun, H., & Zhang, J. (2022). Assessment literacy of college EFL teachers in China: Status quo and mediating factors. *Studies in Educational Evaluation, 74*, 101-117. https://doi.org/10.1016/j.stueduc.2021.101117

Tajeddin, Z., Alemi, M., & Yasaei, H. (2018). Classroom assessment literacy for speaking: Exploring novice and experienced English language teachers' knowledge and practice. *Iranian Journal of Language Teaching Research, 6*(3), 57-77.

Tajeddin, Z., Khatib, M., & Mahdavi, M. (2022). Critical language assessment literacy of EFL teachers: Scale construction and validation. *Language Testing, 2022*, 02655322211057040. https://doi.org/10.1177/02655322211057040

Tao, N. (2014). *Development and validation of classroom assessment literacy scales: English as a foreign language (EFL) instructors in a Cambodian higher education setting* (Publication No. 3685135) [Doctoral dissertation, Victoria University]. ProQuest Dissertations and Theses Global.

Tavassoli, K., & Farhady, H. (2018). Assessment knowledge needs of EFL teachers. *Teaching English Language, 12*(2), 45-65. https://doi.org/10.22132/tel.2018.82201

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/10.1177/0265532213480337

Vitali, G. J. (1994). *Factors influencing teacher's practices in an assessment driven reform* (ERIC No. ED373053).

Walters, F. S. (2010). Cultivating assessment literacy: Standards evaluation through language-test specification reverse engineering. *Language Assessment Quarterly, 7*(4), 317-342. https://doi.org/10.1080/15434303.2010.510208

Wang, J. R., Kao, H. L., & Lin, S. W. (2010). Pre-service teachers' initial conceptions about assessment of science learning: The coherence with their views of learning science. *Teaching and Teacher Education, 26*(3), 522-529. https://doi.org/10.1016/j.tate.2009.06.015

Wang, T., Wang, K., & Huang, S. (2008). Designing a web-based assessment environment for improving pre-service teacher assessment literacy. *Computers and Education, 51*, 448-462. https://doi.org/10.1016/j.compedu.2007.05.012

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

Zamani, R., & Ahangari, S. (2016). Characteristics of an effective English language teacher (EELT) as perceived by learners of English. *International Journal of Foreign Language Teaching and Research*, 4(14), 69-88.