Impact of Language Assessment Literacy Enhancement on Iranian EFL Teachers' Assessment of Learning and Attitude: A Comparative Study of Novice and Experienced Teachers

Afrooz Dorri [1] [iD], Hossein Heidari Tabrizi [1]* [iD], Ahmadreza Lotfi [1] [iD]

[1] Islamic Azad University, Isfahan, Iran

Abstract

This study sought to explore the effect of language assessment literacy enhancement on novice and experienced Iranian English-as-a-Foreign-Language (EFL) teachers' assessment of learning performance and their attitude towards it. To achieve this aim, 80 Iranian EFL teachers with varying teaching experience were selected using convenience sampling and categorized into a novice (n = 40) and an experienced (n = 40) group. After completing a researcher-made Assessment of Learning (AoL) knowledge pre-test, the participants received a Language Assessment Literacy Enhancement (LALE) treatment across 13 instructional sessions, which was followed by an AoL post-test. A semi-structured interview was also administered to 10 of them to examine their attitudes. The results of the independent samples *t*-test showed that experienced teachers outperformed novice teachers in assessment of learning, a finding that can be interpreted within the context of the present study, in light of their greater prior familiarity with summative assessment practices, as well as the formative-oriented focus of the language assessment literacy intervention based on Popham's (2009) framework, in which summative assessment was not emphasized to the same extent. The results also indicated that the participants acknowledged the advantages that recent research has demonstrated for formative assessment, such as motivating learners and making assessments more objective. However, they maintained that, due to teachers' insufficient language assessment knowledge, limited time, and the dominance of summative assessment in Iran's education system, summative assessment remains the practical, accepted, expected, and preferred form of assessment in Iran. The results carry implications for policymakers, teachers, educators, and coursebook developers, among others.

* **Corresponding Author:** Hossein Heidari Tabrizi, Professor of Applied Linguistics, Department of English, Isf. C, Islamic Azad University, Isfahan, Iran, **Email**: heidaritabrizi@iau.ac.ir

1. INTRODUCTION

Assessment is one of the fundamental elements of successful second or foreign language learning; however, its importance is often neglected and often reduced to pen-and-paper tests (Earl, 2003, 2013). However, in recent years, the field of language assessment has witnessed a major shift of attention from restricting language assessment to merely teacher-oriented quantitative mechanisms to more interactive assessments in which the roles of students, teachers, and other stakeholders are acknowledged (Earl, 2013; Lam, 2016, 2020).

This change emphasizes the need to enhance teachers' Language Assessment Literacy (LAL), which encompasses their knowledge, skills, and principles for creating, interpreting, and applying assessment (Fulcher, 2012). Meanwhile, research has revealed that effectively implementing formative assessment in EFL classrooms relies on teachers' assessment literacy (Li & Yongqi Gu, 2023; Schellekens et al., 2023; van der Vleuten & Schellekens, 2024). Recent research revealed that while LAL has gained unprecedented attention in the field of EFL assessment, many teachers still have problems incorporating assessment concepts into their everyday classroom activities (Cui et al., 2025; Gan & Lam, 2020; Giraldo, 2020, 2021; Levi & Inbar-Lourie, 2020; Weng & Shen, 2022) and that its effectiveness is contingent upon several parameters, such as teachers' attitudes towards it (Jannati, 2015; Nasution & Putri, 2025) and their teaching experience (Weng & Shen, 2022; Yan & Fan, 2021).

For example, in Iran, even a brief overview of the extensive related literature shows that although the role and conception of assessment have changed in the field of Teaching English-as-a-Foreign-Language (TEFL), this shift of focus has not been reflected in the TEFL context in Iran (Barootchi & Keshavarz, 2002; Janatifar & Marandi, 2018; Razavipour et al., 2011; Shah Ahmadi & Ketabi, 2019). One of the reasons is that, to be employed as an English teacher in Iran, candidates typically need to demonstrate proficiency in language and teaching abilities. However, they are not evaluated for their LAL. Consequently, the teachers' LAL and its leading role in realizing teaching and learning objectives are usually neglected, leading to Iranian EFL teachers' adopting the dominant assessment type, which is summative assessment (Janatifar & Marandi, 2018; Shah Ahmadi & Ketabi, 2019).

These issues are indicative of the importance of introducing Language Assessment Literacy Enhancement (LALE) programs into teacher training courses (Weng & Shen, 2022). Accordingly, the present study attempts to investigate how LALE interventions influence novice and experienced Iranian EFL teachers' assessment knowledge and their attitudes toward formative assessment in classroom contexts.

2. LITERATURE REVIEW

Although the field of assessment has experienced a paradigm shift, assessment has not been reduced to standardized large-scale exams; rather, it encompasses a broader range of methods and approaches for enhancing learning and evaluating performance (Schmitz et al., 2011). Accordingly, Assessment of Learning (AoL), Assessment for Learning (AfL), and Assessment as Learning (AaL) were distinguished.

The traditional assessment is known as AoL and is mainly used for summative purposes such as certification, measurement, and accountability (Earl, 2003, 2013; Serafini, 2000). AoL focuses on technical aspects like regulations, standards, and efficiency and aims at confirming students' knowledge, evaluating curriculum proficiency, and determining placements or future programs (Western and Northern Canadian Protocol [WNCP], 2006). AoL does not enhance learning processes, but quantifies and evaluates learners' acquired skills and knowledge through

reliable and valid methods (Chong, 2017). AoL often emphasizes examination-based assessments, which are used by governing bodies. Hence, teachers have limited autonomy, and students, viewed as passive test-takers, are expected to prepare for assessments through repetitive exercises and familiarity with high-stakes tests (Lee & Coniam, 2013).

By contrast, AfL focuses on improving student learning by supplying teachers with diagnostic evidence that informs adjustments to instruction (Black & William, 2004). This type of formative assessment involves identifying students' strengths and weaknesses, offering feedback, and helping them understand assessment criteria and learning objectives (Earl, 2003, 2013). Teachers play a central role in AfL because they are regarded as mediators who provide constructive feedback, adapt instructional strategies, and facilitate mutual communication to meet learners' needs (Serafini, 2000). Effective AfL considers teachers' beliefs, learners' readiness, educational settings, and system-level policies (Carless, 2007). The ultimate goal of AfL is to adjust teaching to students' progress through a continuous cycle of observation, feedback, and support. It focuses on enhancing learning through interactive feedback and shared responsibilities within a structured framework (Engeström, 1999).

Finally, AaL centers learners in the assessment process and emphasizes their active role in self-reflection and self-regulation to bridge gaps in their knowledge. It focuses on learners identifying personal learning objectives and enhancing their understanding through formative assessment practices. It requires learners to connect their prior knowledge to current challenges and develop practical solutions. AaL encourages learners to critically assess their progress, set individualized goals, and monitor their learning through self-regulation strategies. Teachers in AaL guide students in assessing both their own and their peers' learning and attempt to foster deeper engagement and personalized instruction (Davison & Leung, 2009).

With the incorporation of AaL and AfL into the assessment field, the concept of assessment literacy has gained increasing attention, prompting researchers to explore and define it. For example, Stiggins (1995) listed the target of assessment, the reason for the assessment, the manner of assessment, the potential difficulties in assessment, and preventive measures in response to these difficulties as the five competencies of proper assessment. However, Fullan (2007) argued that developing assessment literacy involves the ability to make well-informed qualitative judgments that are closely tied to the specific classroom context. DeLuca and Klinger (2010) also argued that assessment literacy refers to teachers' assessment knowledge (for measuring students' learning), and their ability to modify and adapt their teaching methodology accordingly. More recently, Kremmel and Harding (2020) also identified the following nine factors as components of the LAL profile: developing and administering language assessments, assessment in language pedagogy, assessment policy and local practices, personal attitudes and beliefs, research and statistical methods, assessment principles and interpretation, language structure, development, and use, preparation and washback, and rating and scoring.

Along with the above theoretical shifts in the conceptions, functions, and purposes of assessment, empirical studies were also conducted in different countries to shed some light on different aspects of these theoretically novel notions. Hence, a surge of interest was observed in empirical studies investigating AaL, AoL, and AfL. An in-depth review of the related literature revealed that the following themes could be identified in studies conducted on different assessment forms: teachers' characteristics (e.g., their attitude towards AoL, AfL, and AaL; their experience level; their educational level) (Alkharusi et al., 2011; Cheng et al., 2004; Kim et al., 2020; Shahzamani & Tahririan, 2021; Sultana, 2019); learners' characteristics (e.g., their attitude towards different forms of assessment; their motivation for applying them; their required language proficiency for deploying AfL and AaL strategies) (Butler et al., 2021; Homayounzadeh & Razmjoo, 2021; Ortiz & Cuellar, 2018); resources (e.g., assessment instruments; the availability

and appropriateness of coursebooks for implementing formative assessment) (Nasr et al., 2019, 2020); and environmental factors (e.g., classroom size) (Ortiz & Cuellar, 2018).

As for LAL, some studies examined EFL teachers' LAL level (Cumming, 2001; Homayounzadeh & Razmjoo, 2021; Xu & Brown, 2017, to name just a few); other studies investigated factors affecting EFL teachers' LAL (Afshar & Ranjbar, 2021; Crusan et al., 2016; Mansouri et al., 2021; Sultana, 2019, among others), and some other studied the best ways for developing teachers' LAL (Fulcher, 2012; Gan & Lam, 2020; Jin, 2010; Mendoza & Arandia, 2009; Vogt & Tsagari, 2014).

Despite the growing body of research on Language Assessment Literacy (LAL), existing studies have predominantly focused on formative-oriented practices, particularly AfL and AaL, while treating AoL as either a taken-for-granted or peripheral construct. As a result, AoL has remained under-theorized and under-investigated, especially in contexts where summative assessment continues to dominate educational decision-making. Moreover, although teaching experience has been identified as a potentially influential factor in teachers' assessment literacy, empirical evidence comparing novice and experienced teachers' AoL knowledge and performance remains scarce. This gap becomes even more salient in centralized educational systems such as Iran, where high-stakes summative assessment largely determines students' academic trajectories and where teachers receive limited formal training in assessment literacy. Consequently, it remains unclear how Language Assessment Literacy Enhancement (LALE) influences teachers' AoL practices across different experience levels, and how teachers themselves perceive AoL within such assessment-driven contexts. Addressing these gaps, the present study seeks to comparatively examine novice and experienced Iranian EFL teachers' AoL performance and their attitudes toward it following the LALE intervention.

Accordingly, the present study attempts to answer the following research questions:

1. Does language assessment literacy enhancement affect novice and experienced teachers' assessment of learning in Iranian high school EFL classes differently?

2. Does language assessment literacy enhancement affect novice and experienced teachers' attitudes toward assessment of learning in Iranian high school EFL classes?

3. METHODOLOGY AND DESIGN

Research Design

The present study adopted an explanatory multi-method design, integrating both quantitative and qualitative data collection methods to achieve its research objectives. The quantitative component employed a descriptive experimental approach, which examined the effect of LALE on novice and experienced Iranian EFL teachers' Assessment of Learning. The quantitative phase was complemented by interviews as the elicitation technique in the qualitative phase.

Participants

From the initial pool of 200 Iranian high school EFL teachers in Najafabad, Isfahan, Iran identified through convenience sampling, 80 were selected based on predefined inclusion criteria: they were teaching English, had various teaching experiences, aged between 24 and 50, held at least a bachelor's degree in English or a related field, were native speakers of Persian, and agreed to participate voluntarily. These 80 qualified teachers were then divided into two groups according to teaching experience: 40 participants had less than three years of teaching experience (14 male, 26 female), while the remaining 40 had more than three years (17 male, 23 female). Participants

with fewer than three years of experience were categorized as the novice group, whereas those with more than three years were categorized as the experienced group, a distinction commonly adopted in previous teacher education and EFL research (Gatbonton, 2008; Tajeddin et al., 2018).

Instruments

The instruments described below were used to obtain the data for this study.

Test of Assessment Knowledge for and of Learning

Based on the Classroom Assessment Literacy Inventory (CALI), proposed by Mertler and Campbell (2005), a researcher-made test was developed, validated, and administered to measure teachers' knowledge of both AfL and AoL. The full test consists of five scenarios, totaling 35 multiple-choice items, each with one correct answer, designed to assess knowledge in both areas. For the purpose of this study, only the 18 items addressing teachers' knowledge of AoL were analyzed, while the 17 items related to AfL were excluded, as the focus of this research is specifically on AoL. This test was adapted in light of the WNCP (2006) model to capture teachers' knowledge of assessment practices (refer to Salimi and Farsi (2020) for the test). The WNCP model is a curriculum and assessment framework which evaluates teachers' knowledge and their capacity to develop, interpret, and implement assessments in a way consistent with the WNCP curriculum principles. The original test was in English, but was translated into Persian by one of the researchers to minimize the risk of misunderstanding and ensure that language barriers would not interfere with participants' ability to select the correct answers. Participants were given approximately 20 minutes to complete the questionnaire, which was considered sufficient given the number of items and their multiple-choice format. Three experts, including professionals in EFL and language assessment from Islamic Azad University, Isfahan (Khorasgan Branch), were consulted to provide their feedback and recommendations regarding the test to establish its content validity; however, as the expert review was qualitative in nature, a numerical content validity index (CVI) was not calculated. Cronbach's alpha coefficient was calculated to assess the internal consistency of the test, yielding a value of 0.78.

Interview on Attitudes toward Assessment Practices

The study employed a researcher-made semi-structured interview with six questions addressing teachers' positive and negative attitudes toward AoL based on the WNCP (2006) model (see Appendix). The interview questions were formulated through an extensive review of the related literature, combined with the researcher's professional knowledge of language teaching and assessment in Iranian high schools. They were also piloted before being used for the main participants. Piloting helped to gain information regarding required time, the wording and ambiguity of questions, and procedural issues, including the order of questions and the introduction format. A total of 10 Iranian high school EFL teachers participated in the interview phase of the study. Each interview lasted approximately 20-30 minutes on average and was conducted individually. Moreover, to ensure content validity of the interview questions, three associate professors of TEFL from Islamic Azad University, Isfahan (Khorasgan Branch), were invited to comment on them. Based on their feedback, all six items were deemed highly relevant for exploring teachers' attitudes toward assessment practices. The experts' feedback also contributed to enhancing the face validity of the interview questions. The researchers implemented the required adjustments to the instrument to guarantee it would generate the detailed data necessary for a precise thematic analysis.

Data Collection Procedures

Prior to data collection, ethical considerations were carefully observed. Participation in the study was entirely voluntary, and informed consent was obtained from all participants. The teachers were notified of the study's purpose and assured that their responses would remain confidential and solely for research use. The participants were made aware that they could withdraw from the study at any stage without incurring any consequences. To ensure anonymity, no identifying information was collected or reported.

During the quantitative phase, 80 Iranian high school EFL teachers were divided into two groups –novice and experienced– according to whether their English teaching experience was below or above three years.

The Language Assessment Literacy Enhancement (LALE) employed as the treatment for EFL teachers in the present research was based on the 13 themes recommended by Popham (2009). The topics covered in these 13 instructional sessions, together with their definitions, are presented in the table below:

Table 1: Topics of Instructional Courses and their Definitions

Session	Topic of Instructional Course	Definitions
1	Purpose of assessment	Gather evidence to know what students can do, what they know, and how they feel about learning.
2	Reliability	Make sure test results are consistent and repeatable.
3	Validity	Show that the test actually measures what it's supposed to measure.
4	Bias	Remove items that unfairly disadvantage students because of race, gender, or background.
5	Item construction	Write and improve good multiple-choice and open-ended questions.
6	Scoring written responses	Use clear rubrics so scoring is fair and consistent.
7	Performance and portfolio assessment	Judge real work (projects, portfolios, peer/ self-checks), not just quizzes.
8	Formative assessment	Use quick checks and feedback during teaching to guide learning.
9	Attitudes/interests/ values	Find out students' feelings, interests, and values.
10	Interpreting large-scale tests	Understand what big standardized test scores really mean.
11	English language learners and students with disabilities	Adjust tests so English learners and students with disabilities are assessed fairly.
12	Preparing for high-stakes tests	Teach appropriate test skills without unfair coaching or abuse.
13	Accountability tests	Decide if a test is suitable for judging teaching or schools.

One week before applying the treatments, a researcher-made Assessment of Learning (AoL) knowledge pre-test was developed, validated, and administered to determine the participants' AoL knowledge. One week after the pre-test, EFL teachers were given the treatment, that is, the LALE

program for assessment of learning. In the experimental phase, the materials were presented to both the novice and experienced teacher groups during 13 instructional sessions. One week after the treatment, the same AoL knowledge test was administered as a posttest to examine whether participation in the LALE program had improved the teachers' knowledge of assessment of learning.

In the qualitative phase of the study, 10 high school EFL teachers, both novice and experienced, were given interviews eliciting their attitudes toward AoL in their English classes. To obtain their spontaneous responses, participants were not informed of the interview's topic before the interview. However, at the outset of the interview, the aim and expected duration of the interview were explained. Each interview lasted approximately 10 to 15 minutes and was conducted in Persian, the participants' first language, to ensure a clear understanding of the questions and to allow the interviewees to express their views freely and accurately.

LALE Treatment

The LALE program constituted the instructional treatment of the present study. The treatment was designed based on the assessment literacy framework proposed by Popham (2009) and aimed to enhance teachers' knowledge and understanding of AoL principles and practices.

The LALE program was implemented over 13 one-hour online instructional sessions, conducted twice a week over a period of approximately 7 weeks. All sessions were delivered synchronously via the Skype platform by the same instructor, who was an expert in English pedagogy and language assessment, in order to ensure instructional consistency across groups. Also, both novice and experienced teachers received identical instructional content.

The treatment was organized in a stage-based and cumulative manner, with each session focusing on a specific component of assessment literacy. The instructional progression was as follows.

Stage 1: Foundations of Assessment Literacy (Sessions 1-2)

The initial sessions focused on introducing the purposes of classroom assessment, with particular attention to distinguishing between assessment of learning, assessment for learning, and assessment as learning. Key concepts such as the role of assessment in instructional decision-making and accountability were discussed, with an emphasis on the function of AoL in high-stakes educational contexts.

Stage 2: Technical Quality of Assessment (Sessions 3-4)

These sessions addressed the technical qualities of assessment, including different forms of reliability (e.g., test-retest, inter-rater, and internal consistency) and types of validity (content-related, criterion-related, and construct-related validity). Practical examples were provided to illustrate how these concepts apply to classroom-based and summative assessments.

Stage 3: Bias, Fairness, and Ethical Considerations (Session 5)

Session five focused on identifying and minimizing bias in assessment, particularly factors that may unfairly disadvantage learners based on background, language proficiency, or individual differences. Ethical considerations in assessment design and score interpretation were also discussed.

Stage 4: Item Construction and Scoring Procedures (Sessions 6-7)

These sessions concentrated on item construction principles, including the development of multiple-choice and constructed-response items aligned with learning objectives. Teachers were also introduced to scoring written responses, rubric design, and strategies for enhancing scoring objectivity and consistency in summative assessment contexts.

Stage 5: Performance-Based and Portfolio Assessment (session 8)

In this stage, teachers were introduced to performance and portfolio assessment as complementary forms of AoL. The session highlighted how authentic assessment tasks can be used for summative decision-making while maintaining acceptable levels of reliability and validity.

Stage 6: Interpreting and Using Assessment Results (sessions 9-10)

These sessions addressed the interpretation of classroom-based and large-scale assessment results, including understanding score meaning, limitations of test scores, and appropriate use of assessment data for grading and reporting purposes.

Stage 7: Assessment in Special Educational Contexts (Sessions 11-12)

The focus of these sessions was on assessing English language learners and students with disabilities, with emphasis on accommodations and adaptations required to ensure fair summative assessment. Preparation for high-stakes and accountability tests was also discussed.

Stage 8: Consolidation and Review (Session 13)

The final session was devoted to reviewing and consolidating the key assessment literacy concepts covered throughout the program. Participants reflected on how AoL principles could be applied within the constraints of the Iranian high school context.

Data Analysis

Data analysis was conducted using IBM SPSS Statistics version 28. To answer the first question, an independent samples *t*-test was run on the post-test AoL scores (continuous quantitative data) obtained from the researcher-made Assessment Knowledge Test to compare novice and experienced teachers' performance following the LALE intervention. Assumptions of normality and homogeneity of variance were checked and satisfied prior to analysis.

To answer the second question, the interview data were audio-recorded and transcribed verbatim to facilitate qualitative content analysis. The analysis was conducted manually, following a thematic coding procedure consisting of three main stages: open coding, axial coding, and selective coding (Strauss & Corbin, 1998).

In the first stage (open coding), the transcribed interviews were read several times to gain familiarity with the data. Meaningful units related to teachers' perceptions and attitudes toward assessment practices were identified and assigned initial codes (e.g., time-consuming assessment, lack of formal training, dominance of summative assessment).

In the second stage (axial coding), conceptually related open codes were grouped into broader categories and sub-themes by identifying relationships and patterns among them (e.g., practical constraints, institutional barriers, motivational effects of assessment literacy).

In the final stage (selective coding), the major themes that captured teachers' overall attitudes toward assessment of learning were refined and integrated into overarching thematic categories.

To enhance the trustworthiness of the analysis, the coding process was conducted independently by two coders. Inter-coder reliability was assessed using Cohen's Kappa, yielding a coefficient of .86, which indicates substantial agreement between the two coders (Landis & Koch, 1977).

Table 2: Three Phases of Coding of the Attitudes of Iranian EFL Teachers About the Practice of Self- and Peer-Assessment Techniques and Writing Rubrics Instruction in Iranian High-School EFL Classes

Open coding	Axial coding	Selective coding
1. Practice of self and peer assessment is considered a very time-consuming, effortful and challenging process	1. Self and peer-assessment techniques are considered to be non-frequent, time-consuming and challenging by teachers	1. Self and peer-assessment techniques are required to be instructed regularly and frequently to Iranian EFL teachers
2. There were poor and insufficient programs for Iranian EFL teachers	2. Assessment criteria instruction is believed to result in higher levels of motivation, self-confidence, and assessment literacy	
3. The instruction of assessment criteria provided EFL teachers with higher levels of motivation, self-confidence, and assessment literacy		
4. The instruction of assessment criteria (language assessment literacy) is considered the best way to turn a subjective judgement into an objective one		

4. RESULTS

Research Question One

Initially, normality of the data distribution and homogeneity of variance were checked and confirmed. Then, an independent samples *t*-test was run, whose descriptive results are presented in Table 3.

Table 3: Test of Normality of Assessment of Learning Scores for Both Groups

	Kolmogorov-Smirnoff			Shapiro-Wilk		
Assessment of learning	Statistic	df	Sig.	Statistic	df	Sig.
Pre-test (novice teachers)	0.08	40	0.20	0.96	40	0.53
Post-test (novice teachers)	0.12	40	0.20	0.96	40	0.50
Pre-test (experienced teachers)	0.09	40	0.20	0.97	40	0.80
Post-test (experienced teachers)	0.10	40	0.20	0.98	40	0.63

The Kolmogorov-Smirnov and Shapiro-Wilk tests' statistics for the pre- and post-test scores indicated the normal distribution of the data ($p > .05$); therefore, the normality assumption was met, and an independent samples *t*-test can be run on the scores.

The questionnaire was administered to participants prior to the AoL treatment, the LALE program, to ensure the homogeneity of the teachers in terms of familiarity with the AoL concepts and issues. Results of the independent samples *t*-test comparing the pre-test scores of both groups of teachers is presented in Table 4 below.

Table 4: Descriptive Statistics of Assessment of Learning Pre-Test Scores of Both Groups

Group	N	M	SD
Experienced teachers	40	10.70	2.01
Novice teachers	40	11.22	1.90

As shown in Table 4, novice teachers obtained higher scores in the pre-test compared to experienced ones. To examine whether the difference between the two groups reached statistical significance, an independent samples *t*-test was performed. The result of the independent samples *t*-test comparing the pre-scores of both groups is shown in Table 5 below.

Table 5: Independent Samples t-Test Comparing Assessment of Learning Pre-Scores of Both Groups

	Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	.09	.75	1.19	78	.23	.52	.43	-.34	1.39
Equal variances not assumed			1.19	77.73	.23	.52	.43	-.34	1.39

As shown in Table 5, the pre-test scores of experienced (M = 10.70, SD = 2.01) and novice teachers (M = 11.22, SD = 1.09; *t*(78) = 1.19, *p* = .23, two-tailed) did not differ significantly. The result of

the independent samples *t*-test comparing the post-test scores of both groups of teachers is presented in Table 6 below.

Table 6: Descriptive Statistics of Assessment of Learning Post-Test Scores of Both Groups

Group	N	M	SD
Experienced teachers	40	13.62	1.71
Novice teachers	40	12.50	2.42

As shown in Table 6, experienced teachers obtained higher scores in the post-test compared to novice ones. To examine whether the difference between the two groups reached statistical significance, an independent samples *t*-test was performed. The finding of the independent samples *t*-test comparing the post-scores of both groups is shown in Table 7 below.

Table 7: Independent Samples t-Test Comparing Assessment of Learning Post-Scores of Both Groups

	Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	3.63	.04	2.39	78	.01	1.12	.47	.18	2.06
Equal variances not assumed			2.39	70.26	.01	1.12	.47	.18	2.06

Results of the independent samples *t*-test indicated a significant difference between experienced and novice teachers in assessment of learning (M = 12.50, SD = 2.42; *t* (78) = 2.39, *p* = .01, two-tailed), with a mean difference of 1.12 (95% CI [0.18, 2.06]) and a medium-to-large effect size (eta squared = .68).

Research Question Two

Teachers' Attitudes

In general, the following major themes were derived based on the analysis and coding of the interviews with the teachers. These themes, along with some interview samples, are presented below:

Theme 1. Teaching self and peer assessment is a very time-consuming and challenging process

This theme was derived from the analysis of the opinions expressed by nine participants. For example, Zahra (age 31) maintained that:

> As far as I know, most Iranian EFL teachers do not use self and peer assessment in their classrooms. One possible reason is that they themselves do not have enough familiarity with these ideas, as they have not received the related pre- or in-service instruction. Another reason is that they have found it to be a lengthy technique whose implementation needs a great amount of time, effort, and energy, which devotes a big part of class time to itself, as the EFL learners are usually reluctant to use and implement self and peer-assessment techniques […]. In fact, in Iran's high school English classrooms, summative assessment is more practical and accepted by students, their parents, and school principals. (Excerpt 1).

Similarly, this theme emerged from the analysis of the views expressed by nine participants. In this regard, Elham (age 28) stated that:

> Instruction, development, and conducting effective self-assessment and peer-assessment strategies and rubric instruction in the EFL courses take a great deal of effort and class time. As these techniques and activities are almost completely new to both teachers and EFL learners in Iranian EFL contexts, they require a great amount of energy, concentration, trial, and effort to be successfully implemented, and this may cause the EFL teachers to be doubtful about using self and peer assessment strategies in their classes. To be honest, I, as an Iranian English teacher, still prefer to use traditional assessment methods since self and peer-assessment, in my opinion, are very challenging, time-consuming, and tiring. (Excerpt 2).

Theme 2. Iranian EFL teachers receive very little or no formal instruction on assessment

This theme was derived from the analysis of the opinions expressed by seven participants. For example, Shadi (age 37) maintained that:

> To my best knowledge, Iranian EFL teachers receive very little or no formal instruction on the assessment and assessment criteria, which is the essential prerequisite for a reliable and valid assessment by the teacher and for the instruction of peer and self-assessment techniques to the EFL learners. In the pre-service and in-service programs, the EFL teachers should be provided with the necessary training courses that enable them to execute self and peer assessment techniques in the class. But unfortunately, I think, most Iranian EFL teachers are dissatisfied with the poor quality of the in-service and pre-service training system. (Excerpt 3)

Similarly, this theme emerged from the analysis of the views expressed by seven participants. In this regard, Reza (age 39) stated that:

> I have received my training mainly through limited writing courses in the university, but mainly through self-study and writing textbooks. Assessment for learning and as

learning are advantageous but because we are not trained to apply them in our classrooms, they are not practical. (Excerpt 4)

Theme 3. The instruction of assessment criteria to EFL teachers and learners provides both of them with higher levels of motivation and writing assessment literacy

This theme was derived from the analysis of the opinions expressed by eight participants. For example, Nasrin (age 42) maintained that:

> I think that after being instructed on the assessment criteria, EFL students obtained enough required knowledge and, consequently, sufficient motivation and self-confidence to go through the processes of self and peer assessment. (Excerpt 5)

Similarly, this theme emerged from the analysis of the views expressed by eight participants. In this regard, Hamid (age 31) stated that:

> Before the EFL students are trained in the what and how of the writing assessment criteria, they are faced with a kind of confusion as they do not know why they have obtained a specific score or how they can analyze, evaluate, and grade their own or their peers' writings. This confusion will, of course, lead to disappointment and loss of motivation for further learning. Being provided with this kind of instruction, on the other hand, opens a new horizon for students to the writing skill and its accurate assessment**. (Excerpt 6)

Theme 4. The instruction of assessment criteria (language assessment literacy) is the best way to turn a subjective judgment into an objective one

This theme was derived from the analysis of the opinions expressed by seven participants. For example, Fatemeh (age 35) maintained that:

> We, as EFL teachers, know that writing skill assessment has a great potential to be subjective if it is done without a standard scoring criterion. Using such a criterion, however, can turn a completely subjective task into an objective one. Intra and inter-rater reliability can be guaranteed by such kind of instruction. To be honest, before this instructional course, when assessing my students' writing tasks, I fluctuated on putting the focus on the content of writing, grammar, punctuation, etc., and I think this may have changed from student to student, as some of them were very good at grammar, some at managing the content, and some in other writing aspects. Teaching assessment criteria to EFL students can also make them critical, objective, and conscious assessors of their own and their peers' writings. But I have gotten used to the summative assessment after years of teaching, and changing my assessment method is difficult for me. If I change my assessment method, children might not accept it because they have also gotten used to summative assessment. (Excerpt 7)

5. DISCUSSION

Previous research has highlighted that teachers' Language Assessment Literacy (LAL) can vary according to their teaching experience, with experienced teachers often demonstrating higher assessment competence than novices (Sultana, 2019; Van der Vleuten & Schellekens, 2024). The present study aimed to examine the effect of LALE on teachers' Assessment of Learning (AoL)

and to compare outcomes between experienced and novice English language teachers. Specifically, the first research question investigated whether experienced and novice teachers differed in their AoL after participating in LALE. The analysis revealed that the difference in post-test scores from the researcher-made Assessment Knowledge Test was statistically significant, with experienced teachers outperforming their novice counterparts. This finding corroborates the findings of previous studies (Edwards, 2017; Gatbonton, 2008), revealing that EFL teachers with more teaching experience practice more summative assessment strategies in their classrooms. In this regard, Tajeddin et al. (2018) found that experienced Iranian EFL teachers' summative assessment knowledge was comparably more than their novice counterparts and more consistency was observed between their knowledge and actual classroom practice. While the findings of this study are broadly consistent with those of Tajeddin et al. (2018), the two studies differ in important respects. Tajeddin et al. (2018) primarily examined teachers' classroom assessment literacy for speaking and focused on the alignment between teachers' knowledge and actual assessment practices. In contrast, this study specifically evaluated the effect of a structured LALE intervention on teachers' AoL knowledge, as measured through a standardized post-test. Therefore, although both studies reported superior performance by experienced teachers, the current findings extend previous research by demonstrating how targeted assessment literacy instruction interacts with teaching experience to influence AoL knowledge outcomes. The results of Shah Ahmadi and Ketabi (2019), in this connection, also showed that although experienced teachers' knowledge of summative assessment (AoL) was sufficient and satisfactory, all teachers with various education and experience levels (novice vs. experienced) had limited and insufficient knowledge about the theoretical and practical dimensions of AfL-AaL (or formative assessment).

The observed difference between novice and experienced teachers can be explicated by the argument that EFL teachers with limited teaching experience are more eager to implement what they have learned in pre-service or in-service programs in their classrooms, compared to their experienced counterparts who are accustomed to summative strategies (Pilcher, 2001; Xu & Brown, 2017). In this regard, Mohammadi (2020) also argued that, since experienced Iranian EFL teachers are affected by job burnout and resistance towards change, novice teachers perform better than them in integrating their language assessment literacy into their actual classroom practices, while experienced teachers have more summative assessment (or AoL) knowledge.

Additionally, the observed difference may partly reflect experienced teachers' prior familiarity with AoL strategies due to years of teaching experience, although this was not directly measured in the present study. Given that AoL strategies are widely practiced in Iran and dominate Iran's language assessment system (Barootchi & Keshavarz, 2002; Janatifar & Marandi, 2018; Razavipour et al., 2011), experienced teachers are often more familiar with them, resulting in their better performance compared to novice teachers who are still becoming familiar with these aspects. In this connection, the findings of Razavipour et al. (2011) showed that more than one-third of experienced Iranian EFL teachers could not provide a correct definition for one of the core concepts of assessment, namely; "reliability," but, due to the dominance of summative assessment in Iranian educational system, they were familiar with it and faced almost no problem in implementing it.

Furthermore, to account for the obtained results, we can refer to Watmani et al.'s (2020) observation that Iranian EFL teachers' language assessment courses fail to provide the required theoretical LAL and how it can be implemented in their classrooms. Accordingly, they possess only rudimentary AfL-AaL knowledge, and their LAL is restricted to summative assessment, which they have acquired after being employed and through years of teaching and assessment experience. Barootchi & Keshavarz (2002) also showed that most Iranian EFL teachers have not

received instruction regarding the theoretical and practical dimensions of LAL in their pre-service and in-service teacher training programs. Thus, they are used to summative assessment and have a superficial and rudimentary knowledge of the fundamental concepts of formative assessment. Consequently, through being involved in Iran's educational system, Iranian EFL teachers gradually learn AoL strategies, which are widely accepted within this context, and, accordingly, as their teaching experience grows, their knowledge and application of AoL strategies progressively increase.

Janatifar and Marandi (2018), also examining the components of LAL in the Iranian EFL context, found that the major components of LAL in Iran are as follows: test design and development, large-scale standardized testing and classroom assessment, beyond-the-test aspects (mostly including social and ethical features of LA), and reliability and validity. As can be seen, AfL-AaL strategies are not among the components of LAL in Iran. Hence, Iranian EFL teachers do not have enough LAL, and they learn only summative assessment strategies, which are widely employed in Iran through teaching.

Finally, experienced teachers' better performance in AoL can also be ascribed to the nature of the instruction material provided to the teachers of the experimental group, as the LALE treatment. The in-depth analysis of Popham (2009) proposed themes, which were adopted in this study for enhancing teachers' assessment literacy, revealing that these themes are mostly concerned with AfL goals, and AoL is not emphasized to the same degree. Accordingly, experienced teachers, having had more opportunities to practice and refine their application of AoL techniques over time, are accustomed to integrating AoL strategies into their existing practices. This consistent practice leads to their better performance in AoL.

As for the second research question examining Iranian EFL teachers' attitudes toward Aol, the qualitative findings revealed that although teachers acknowledged the motivational and objectivity-enhancing advantages of AfL and AaL, they considered these approaches impractical in the Iranian context due to limited classroom time, insufficient formative assessment literacy, and the dominance of summative assessment as the accepted and expected form of evaluation. Overall, these results are broadly consistent with prior studies conducted in the Iranian EFL context (Janatifar & Marandi, 2018; Razavipour et al., 2011; Shah Ahmadi & Ketabi, 2019), which similarly reported that despite teachers' positive attitudes toward formative assessment, summative assessment continues to prevail because of institutional expectations and systemic constraints.

The present findings also align with Nasr et al. (2019, 2020), who reported that Iranian EFL teachers perceived formative assessment practices as desirable yet difficult to implement due to contextual barriers, such as large class sizes, time pressure, and assessment-driven curricula. Likewise, Homayounzadeh and Razmjoo (2021) reported that both teachers and learners in Iran tend to favor summative assessment, as it is perceived as more transparent, familiar, and aligned with high-stakes decision-making processes.

However, unlike some recent international studies (e.g., Li & Yongqi Gu, 2023; Schellekens et al., 2023), which demonstrated that sustained professional development and institutional support can gradually shift teachers' assessment practices toward formative orientations, the present study highlights that such structural support is largely absent in the Iranian educational system. This lack of systematic support appears to intensify teachers' resistance to moving away from summative assessment, particularly among experienced teachers who have long relied on traditional assessment practices.

Furthermore, the finding that experienced teachers expressed greater difficulty in transitioning to formative assessment contrasts with studies conducted in less exam-driven contexts (e.g., Yan & Fan, 2021), where experienced teachers were more flexible in adapting

assessment innovations. This discrepancy underscores the context-dependent nature of teachers' assessment attitudes and suggests that entrenched summative assessment cultures, such as that of Iran, may override individual teachers' willingness to adopt formative approaches.

Overall, the qualitative results of this study extend earlier research by demonstrating that while Iranian EFL teachers conceptually value formative assessment, meaningful change requires not only increased language assessment literacy but also systemic reforms at the curricular, institutional, and policy levels.

This argument accounts for and explicates the observed attitudes of the participating Iranian EFL teachers, arguing that although formative assessment is advantageous, summative assessment is what is accepted, practical, and expected in Iran's educational system.

6. CONCLUSIONS

The present study examined the effect of LALE on novice and experienced Iranian EFL teachers' AoL performance and their attitude towards AoL. The results revealed that experienced teachers outperformed their novice counterparts for the following reasons: the dominance of summative assessment in Iran, limited attention to formative assessment issues in Iran's teacher training courses, and the focus of the adopted LALE program on AfL. Regarding Iranian teachers' attitude towards AoL, results indicated that participants held a favorable view of formative assessment and regarded it as more motivational and objective. However, they argued that, due to some factors, such as time limitations, teachers' limited LALE, and the dominance of summative assessment in Iran's educational system, summative assessment is the practical, accepted, and expected assessment form in Iran. Hence, we can conclude that, in some countries, such as Iran, these factors lead to the dominance of AoL, despite the results of the burgeoning research in this field ascertaining the advantages of formative assessment.

The findings of this study yield important implications for policymakers to make LAL a mandatory requirement for teacher employment and incorporate LAL training in pre-service and in-service courses to ensure that teachers are equipped to use assessment as a tool for enhancing learning rather than merely for evaluation purposes. It also guides the development of national curricula, which allocate more time to foreign language instruction to ensure the effective application of AfL-AaL strategies. The findings can also help coursebook developers to include more activities based on AfL and AaL fundamental concepts and teacher training book developers to incorporate more information about the theoretical aspects of AfL-AaL strategies, their advantages, and the best ways of applying them in their classrooms. Finally, given that the present research included both novice and experienced teachers, its findings can be used for performing targeted interventions and designing specific professional development programs that respond to the specific needs of each group.

This study is not without its limitations. First, teachers' performance regarding LAL was assessed and evaluated using self-report questionnaires and interviews. Hence, there might be inconsistencies between participants' reported knowledge and their actual classroom practice. Second, the role of some factors, such as classroom size, learners' age, or proficiency levels, was not taken into account. However, they might affect teachers' attitudes toward the implementation of formative assessment strategies and, accordingly, affect the research findings. Third, participants' prior knowledge of AoL strategies was not directly measured or controlled. Consequently, differences between novice and experienced teachers may partly reflect prior experience with these strategies, which must be considered in the interpretation of the results.

Further research could extend the present study by examining the impact of language assessment literacy enhancement in diverse educational settings and among teachers of different

subject areas or levels of experience. Additionally, subsequent studies might investigate the role of factors not controlled here, such as teachers' prior knowledge of assessment strategies, classroom characteristics, or student proficiency, to better understand how these variables interact with professional development programs. Longitudinal studies could also provide insights into the sustained effects of such interventions on teachers' assessment practices.

References

Afshar, H. S., & Ranjbar, N. (2021). EAP teachers' assessment literacy: From theory to practice. *Studies in Educational Evaluation, 70*(1), 101042. https://doi.org/10.1016/j.stueduc.2021.101042

Alkharusi, H., Kazem, A. M., & Al-Musawai, A. (2011). Knowledge, skills, and attitudes of preservice and inservice teachers in educational measurement. *Asia-Pacific Journal of Teacher Education, 39*(2), 113–123. https://doi.org/10.1080/1359866X.2011.560649

Barootchi, N., & Keshavarz, M. H. (2002). Assessment of achievement through portfolios and teacher-made tests. *Educational Research*, *44*(3), 279–288. https://doi.org/10.1080/00131880210135313

Black, P., & William, D. (2004). The formative purpose: Assessment must first promote learning. In M. Wilson (Ed.) *Towards coherence between classroom assessment and accountability: 103rd Yearbook of the National Society for the Study of Education* (pp. 20–50) University of Chicago Press.

Butler, Y. G., Peng, X., & Lee, J. (2021). Young learners' voices: Towards a learner-centered approach to understanding language assessment literacy. *Language Testing, 38*(3)*,* 429–455. https://doi.org/10.1177/0265532221992274

Carless, D. (2007). Conceptualizing pre-emptive formative assessment. *Assessment in Education: Principles, Policy and Practice, 14*(2), 171–184. https://doi.org/10.1080/09695940701478412

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

Chong, I. (2017). Assessment dialogues between teachers and students using e-writing portfolios. *TESOL Journal, 8*(1). 240–243. https://doi.org/10.1002/tesj.312

Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing, 28*(1), 43–56. https://doi.org/10.1016/j.asw.2016.03.001

Cui, Y., Liu, Y., & Yu, H. (2025). Enhancing English teachers' language assessment literacy through a professional learning community: A collaborative autoethnography study. *System, 134*(98), 103799. https://doi.org/10.1016/j.system.2025.103799

Cumming, A. (2001). ESL/EFL instructors' practices for writing assessment: Specific purposes or general purposes? *Language Testing, 18*(2), 207–224. https://doi.org/10.1177/026553220101800206

Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly*, *43*(3), 393–415. https://doi.org/10.1002/j.1545-7249.2009.tb00242.x

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

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks.

Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning*. Corwin Press.

Edwards, F. (2017). A rubric to track the development of secondary pre-service and novice teachers' summative assessment literacy. *Assessment in Education: Principles, Policy & Practice*, *24*(2), 205–227. https://doi.org/10.1080/0969594X.2016.1245651

Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R.L. Punamäki (Eds.), *Perspectives on activity theory* (pp. 19–38). Cambridge University Press. https://doi.org/10.1017/CBO9780511812774.003

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

Fullan, M. (2007). *The new meaning of educational change*. Teachers College Press.

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

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

Giraldo, F. (2020). A post-positivist and interpretive approach to researching teachers' language assessment literacy. *Profile: Issues in Teachers' Professional Development*, *22*(1), 189–200. https://doi.org/10.15446/profile.v22n1.78188

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

Homayounzadeh, Z., & Razmjoo, S. A. (2021). Examining 'assessment literacy in practice' in an Iranian context: Does it differ for instructors and learners? *Teaching English as a Second Language Quarterly, 40*(2), 1–45. https://doi.org/10.22099/jtls.2021.40269.2978

Jannati, S. (2015). ELT teachers' language assessment literacy: Perceptions and practices. *The International Journal of Research in Teacher Education*, *6*(2), 26–37. https://ijrte.inased.org/makale/152

Janatifar. M., & Marandi, S. S. (2018). Iranian EFL teachers' language assessment literacy (LAL) under an assessing lens. *Applied Research on English Language*, *7*(3), 361–382. https://doi.org/10.22108/are.2018.108818.1221

Jin, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing, 27*(4), 555–584. https://doi.org/10.1177/0265532209351431

Kim, A. A., Chapman, M., Kondo, A., & Wilmes, C. (2020). Examining the assessment literacy required for interpreting score reports: A focus on educators of K–12 English learners. *Language Testing, 37*(1)*,* 54–75. https://doi.org/10.1177/0265532219859881

Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the language assessment literacy survey. *Language Assessment Quarterly*, *17*(1), 100–120. https://doi.org/10.1080/15434303.2019.1674855

Lam, R. (2016). Assessment as learning: Examining a cycle of teaching, learning, and assessment of writing in the portfolio-based classroom. *Studies in Higher Education*, *41*(11), 1900–1917. https://doi.org/10.1080/03075079.2014.999317

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.2019.01.006

Landis, J. R. & Koch, G. G. (1977). *The measurement of observer agreement for categorical data.* Biometrics, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lee, I., & Coniam, D. (2013). Introducing assessment for learning for EFL writing in an assessment of learning examination-driven system in Hong Kong. *Journal of Second Language Writing, 22*(1), 34–50. https://doi.org/10.1016/j.jslw.2012.11.003

Li, J., & Yongqi Gu, P. (2023). Developing classroom-based formative assessment literacy: An EFL teacher's journey. *Chinese Journal of Applied Linguistics*, *46*(2), 198–218. https://doi.org/10.1515/CJAL-2023-0204

Levi, T., & Inbar-Lourie, O. (2020). Assessment literacy or language assessment literacy: Learning from the teachers. *Language Assessment Quarterly*, *17*(2), 168–182. https://doi.org/10.1080/15434303.2019.1692347

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

Mansouri, B., Molana, K., & Nazari, M. (2021). The interconnection between second language teachers' language assessment literacy and professional agency: The mediating role of institutional policies. *System, 103*, 102674. https://doi.org/10.1016/j.system.2021.102674

Mendoza, A. A. L., & Arandia, R. B. (2009). Language testing in Colombia: A call for more teacher education and teacher training in language assessment. *Profile: Issues in Teachers' Professional Development, 11*(2), 55–70. https://doi.org/10.15446/profile.v11n2.4662

Mertler, C. A., & Campbell, C. (2005). *Measuring teachers' knowledge & application of classroom assessment concepts: Development of the Assessment Literacy Inventory. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada. https://eric.ed.gov/?id=ED490355*.

Mohammadi, A. (2020). *A mixed-methods study on the teacher assessment literacy of ELT instructors versus content instructors (Unpublished doctoral dissertation). Islamic Azad University.*

Nasr, M., & Bagheri, M. S., & Sadighi, F. (2020). Iranian English language teachers' perceptions of monitoring and scaffolding practices of assessment for learning: A focus on gender and class size. *International Journal of Foreign Language Teaching and Research, 8*(29), 75–90. https://sanad.iau.ir/Journal/jfl/Article/990688

Nasr, M., Bagheri, M. S., Sadighi, F., Rassaei, E., & van de Weijer, J. (2019). Iranian EFL teachers' assessment for learning practices and barriers: Do textbooks taught and teaching context matter? *Cogent Arts & Humanities*, *6*(1), 1646691. https://doi.org/10.1080/23311983.2019.1646691

Nasution, P. T., & Putri, H. (2025). Graduate students' perceptions and attitudes of peer assessment on academic writing. AMI: *Jurnal Pendidikan Dan Riset, 3*(2), 6–20. http://jurnaltarbiyah.uinsu.ac.id/index.php/ami

Ortiz, S. M. R., & Cuéllar, M. T. A. (2018). Authentic tasks to foster oral production among English as foreign language learners. *HOW*, *25*(1), 51–68. https://doi.org/10.19183/how.25.1.362

Othman, A., Osman, K., & Othman, N. (2024). Assessment value: A systematic literature review on assessment as, for and of learning in school. *International Journal of Academic Research in Progressive Education and Development, 13*(1), 995–1013. https://doi.org/10.6007/ijarped/v13-i1/20725

Pilcher, J. K. (2001). Pilcher, J. K. (2001, March 2). *The standards and integrating instructional and assessment practices*. Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education, Dallas, TX. https://files.eric.ed.gov/fulltext/ED451190.pdf

Popham, W. J. (2009). Assessment literacy for teachers: faddish or fundamental? *Theory into Practice, 48*(1), 4–11. https://doi.org/10.1080/00405840802577536

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

Salimi, E. A., & Farsi, M. (2020). Probing into EFL teachers' assessment literacy and teaching experience: The case of native ESL and non-native EFL teachers. *Journal of English Language Teaching and Learning*, *12*(25), 269–288. https://doi.org/10.22034/elt.2020.10684

Schellekens, L. H., Kremer, W. D. J., Van der Schaaf, M. F., Van der Vleuten, C. P. M., & Bok, H. G. J. (2023). Between theory and practice: Educators' perceptions on assessment quality criteria and its impact on student learning. *Frontiers in Education, 8*(1), 1147213. https://doi.org/10.3389/feduc.2023.1147213

Schmitz, B., Klug, J., & Schmidt, M. (2011). Assessing self-regulated learning using diary measures with university students. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 251–266). Routledge.

Serafini, F. (2000). Three paradigms of assessment: Measurement, procedure, and inquiry, *The Reading Teacher, 54*(4), 384–393. https://www.jstor.org/stable/20204924

Shah Ahmadi, M. R., & Ketabi, S. (2019). Features of language assessment literacy in Iranian English language teachers' perceptions and practices. *Teaching English as a Second Language Quarterly, 38*(1), 191–223. https://doi.org/10.22099/jtls.2020.34843.2739

Shahzamani, M., & Tahririan, M. H. (2021). Iranian medical ESP practitioners' reading comprehension assessment literacy: Perceptions and practices. *Iranian Journal of English for Academic Purposes*, *10*(1), 1–15. https://doi.org/20.1001.1.24763187.2021.10.1.1.5

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappa*, *77*(3), 238–245. https://eric.ed.gov/?id=EJ514726

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.)*.* Sage Publications.

Sultana, N. (2019). Language assessment literacy: An uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, *9*(1), 1–14. https://doi.org/10.1186/s40468-019-0077-8

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. https://doi.org/10.30466/ijltr.2018.120601

Van der Vleuten, C., & Schellekens, L. (2024). Facilitating programme-level change in higher education. In C. Evans & M. Waring (Eds.), *Research handbook on innovations in assessment and feedback in higher education*: *Implications for teaching and learning* (pp. 443–454). Edward Elgar Publishing. https://doi.org/10.4337/9781800881600.00033

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

Watmani, R., Asadollahfam, H., & Behin, B. (2020). Demystifying language assessment literacy among high school teachers of English as a foreign language in Iran: Implications for teacher education reforms. *International Journal of Language Testing*, *10(*2), 129–144. https://www.ijlt.ir/article_118025.html

Weng, F., & Shen, B. (2022). Language assessment literacy of teachers. *Frontiers in Psychology*, *13*(1), 864582. https://doi.org/10.3389/fpsyg.2022.864582

Earl, L. and Katz, S. (2006). *Rethinking classroom assessment with purpose in mind: Assessment for learning, assessment as learning, assessment of learning*. Manitoba Education, Citizenship and Youth. https://www.edu.gov.mb.ca/k12/assess/wncp/full_doc.pdf

Xu, Y., & Brown, G. T. (2017). University English teacher assessment literacy: A survey-test report from China. *Papers in Language Testing and Assessment*, *6*(1*)*, 133–158. http://hdl.handle.net/10722/245352

Yan, X., & Fan, J. (2021). "Am I qualified to be a language tester?": Understanding the development of language assessment literacy across three stakeholder groups. *Language Testing*, *38*(2), 219–246. https://doi.org/10.1177/0265532220929924

Appendices

Appendix A: Interview on Teachers' Attitudes toward Assessment Practices (Researcher-made)

The interview questions were originally developed in English and were translated into Persian for the interviews. The English version is provided here for reference purposes.

1. In your opinion, what are the positive points and strengths of assessment practices of learning in your English class?
2. According to you, what are the negative points and weaknesses of assessment practices of learning in your English class?
3. What do you believe are the primary purposes of assessment of learning in Iranian EFL classrooms?
4. What types of assessment of learning methods (e.g., standardized tests, quizzes, oral exams) do you use most frequently to assess your students' learning? Why?
5. How do you decide which assessment of learning method to use in different classrooms?
   Have you received any formal training on designing and implementing effective assessment of learning methods? If so, how helpful was it?