The Impact of Language Assessment Literacy Enhancement (LALE) on Iranian High School EFL Students' Knowledge of Assessment as Learning in Writing

Afrooz Dorri[1], Hossein Heidari Tabrizi[2]*, Ahmadreza Lotfi[3]

ARTICLE INFO	ABSTRACT
	The impact of self-and peer-assessment strategies (assessment as learning) on the development of writing skill of EFL learners is a research area requiring more research studies. The present study aimed to investigate the impact of language assessment literacy enhancement (LALE) on Iranian high school English as Foreign Language (EFL) students' assessment as learning of writing skill. It also aimed to examine if LALE affects Iranian EFL students' attitudes toward assessment as learning. To this end, 80 intermediate-level high-school EFL learners were selected and randomly categorized into the experimental and control groups. Both groups wrote an essay in the pre-test phase. Then, the English as Second Language (ESL) Composition Scale was used to teach the students what good writing is and what criteria and standards they are supposed to learn to be able to write and rate the essays of their own and their peers. The control group received no instructional information on assessment rubrics. The experimental and control groups wrote another essay on a specific topic in the post-test phase. In the qualitative phase of the study, ten high school EFL students from the experimental group were interviewed regarding their attitudes toward the practice of assessment as learning of the writing skill in their English classes. The findings indicated that LALE significantly affected Iranian high school EFL students' assessment as learning of writing skill. Moreover, students believed peer and self-assessment techniques are rarely implemented in Iranian high school EFL classes. They were also uncertain and felt uncomfortable judging, evaluating, criticizing, and rating their peers. Besides, they felt that they were not knowledgeable and capable enough to play the role of an assessor.

1. Introduction

Assessment, in all educational contexts, has significant importance (Naeini, 2013). As stated by Crooks (2001), assessment refers to any process providing the teacher with the required information about students' achievement, progress, or thinking.

In line with paradigmatic developments in assessment, Earl (2013) theorized the concept of "assessment as learning" (AaL), in which learners are situated at the assessment process center. In AaL, learners, rather than teachers, make connections between assessment and learning because they are

---

[1] Department of English, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran, Email: afroozdorri@gmail. com
[2] Department of English, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran, Email: heidaritabrizi@gmail. com
[3] Department of English, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran, Email: lotfi.ahmadreza@gmail.com

different constructs, although Assessment for Learning (AfL) is essentially a superset of AaL. As opposed to the summative, retrospective nature of Assessment of Learning (AoL), AaL is committed to capacitating learners to engage in self-reflection and self-regulation to define individual learning objectives and fill in their zone of proximal development (Vygotsky, 1978) by bridging the gap between their current and perspective knowledge.

Recently, the notion of assessment literacy has gained considerable interest and, thus, attracted scholarly attempts to define the construct. Fullan (2007) maintained that gaining assessment literacy includes a growing high specialty in teachers, which, as Sadler (1989) posited, encompasses the capability of making "sound qualitative judgments" closely associated with the context of the respective classroom. Davies (2008) indicated knowledge, skills, and principles as the building blocks of assessment literacy. As maintained by Fulcher (2012), assessment literacy has three interdependent dimensions: sufficient skills and knowledge for the design and evaluation of a test, acquaintance with the ethical considerations, and social, political, and cultural contextualization of assessment to examine its consequential validity in society and institutions.

2. Review of Literature

Language assessment is considered the basic and inseparable part of every instructional curriculum. Assessment as learning (AaL) accentuates the role of learners as the main agents to facilitate their learning through a process of metacognition (WNCP, 2006). Learners should be engaged and instructed on how to assess their own and their peers' learning, leading to better delivery of instruction for individualistic learning by teachers and learners (Effendi, 2020). In AaL, learners are also "active agents," making real connections between their learning and their current assessment-induced performance (Chong, 2017).

Several studies have evaluated the relationship between EFL learners' language assessment literacy enhancement and their knowledge of assessment as learning operationalized as writing skill. Fahimi and Rahimi (2015) aimed to investigate the influence of incorporating self-assessment practices on improving the writing skill of Iranian EFL learners. In the pre-test phase, students had no idea how to assess and mark their writing, and the teacher did the assessment. In the treatment phase, the learners were instructed on the required criteria for assessing writing. During these sessions, learners wrote and marked four pieces of writing based on the instructed criteria, and the teacher corrected their writing, giving her marks. A self-assessment questionnaire was administered before and after the treatment, and learners' and teachers' scoring of writing tasks was used to collect the required data. The obtained results demonstrated the gradual development of the EFL learners' writing skill during the treatment period. The interviews with some of the learners and the teacher also showed the positive attitudes of the teacher and learners toward self-assessment.

Nemati and Ghafoori (2016) investigated the influence of assessment types on the L2 writing development of Iranian EFL learners. The study involved a pre-test of writing, an intervention, and a post-test of writing. The findings showed that different assessment types, including teacher, self, and peer assessment, led to the increase of acquaintance of learners with the writing models, wording, correctness, grammar, dictation, pragmatics, etc. The findings showed that these kinds of activities can improve learners' writing. Results analysis showed that different assessment types, like teacher assessment, peer assessment, and self-assessment, significantly developed the writing skill of Iranian EFL learners. The results also indicated that self-assessment was the pioneer assessment type, and peer assessment, compared to teacher assessment, ended in more fruitful results.

Ndoye (2017) investigated students' perceptions of the processes and mechanisms through which self- and peer-assessment can help to engage them in and enhance their learning. The results demonstrated that based on students' perceptions, peer and self-assessment could significantly help to improve their learning through learners' collaboration, effective feedback, and a supportive learning environment. A higher level of awareness of course requirements and the ability to recognize the learning gaps and acquire the required strategies to fill them were the mechanisms through which students perceived that self- and peer-assessment could improve their sense of responsibility towards their learning. Students' dispositions to work in groups can affect the advantages of self-and peer assessment.

Fathi et al. (2021) evaluated the influence of conducting self- and peer-assessment tasks on the writing self-efficacy and writing performance of Iranian EFL learners. Findings revealed that both self- and peer-assessment tasks significantly led to the improvement of the participants' writing performance and writing self-efficacy.

A study by Aslanoglu (2022) investigated the effect of peer- and self-assessment practices on the 9th-grade students' writing skills. The findings showed a significant difference between the pre-test and post-test scores of experimental groups whose members participated in peer- and self-assessment tasks, while no significant difference was found between the pre and post-test scores of the control group. Hence, peer and self-assessment tasks were influential in improving students' writing skills and their interests and attitudes toward writing.

Chen (2023) investigated the effectiveness of short and extended assessment literacy-building interventions to promote the development of students' assessment literacy and subject learning to achieve AaL. The findings revealed that both the short and extended assessment literacy-building interventions could enable learners to go through positive changes in their perceptions of assessment. In addition, there was a significant correlation between students' perceptions of assessment and students' assessment literacy, while there was no significant difference between short and extended assessment literacy interventions regarding promoting positive changes to students' perceptions of assessment, developing students' assessment literacy, and enhancing subject learning.

Mohammadi et al. (2024) evaluated learners' assessment of their writings before and after receiving the rating criteria. The findings showed a significant difference between the students' assessment before and after being provided with the assessment criteria and the practice session. In the same line, they were not familiar with the self-assessment procedure. After receiving the teachers' agreed-upon assessment criteria, the students got familiar with different components of writing. Moreover, they learned how to apply those criteria in their assessment.

He and Wang (2024) assessed the implementation of assessment as learning (AaL) in a writing course at a university in China and found that AaL improved students' learning by offering timely focused feedback, creating personalized learning experiences, and encouraging learner engagement. Hedayati and Khoorsand (2024) explored the impact of assessment as learning on the writing development of intermediate EFL learners, and the results confirmed the participants' significant improvement in writing skills. Furthermore, the interviews with participants revealed that AaL was a new and good experience; they felt less stress in the writing class and enhanced their classroom participation and engagement.

As the review of the related literature reveals, the majority of the research studies conducted in the domain of language assessment literacy heavily rely on EFL teachers, while EFL learners can also make great use of assessment as a learning tool. The language assessment literacy of EFL learners and the effect of language assessment literacy enhancement on the EFL learners' knowledge of assessment of, for, and as learning is underexplored. The influence of EFL learners' language assessment literacy on their knowledge of assessment as learning can be considered the most overlooked component. Few studies examined the role and importance of language assessment literacy enhancement (LALE) in developing the EFL learners' knowledge of assessment as learning. It is missing among the research studies and needs to be considered an important research subject in the future. Therefore, this study aimed to investigate if language assessment literacy enhancement (LALE) can significantly affect Iranian high school EFL students' assessment as learning, operationalized as writing skill. It also aimed to examine which writing components are improved due to language assessment literacy enhancement. Another aim of the present study was to examine if language assessment literacy enhancement (LALE) affects Iranian high school EFL students' attitudes toward assessment as learning. The following questions were posed for the current study:

1. Does language assessment literacy enhancement (LALE) significantly affect students' assessment as learning, operationalized as writing skill, in Iranian high school EFL classes compared with the control group?

2. Which of the writing skills (i.e., grammar, vocabulary, mechanics, organization, and content) was more significantly influenced by language assessment literacy enhancement (LALE)?

3. What are the Iranian high school EFL students' attitudes toward language assessment literacy enhancement (LALE) in classes?

3. Method

3.1. Design

The present study used a mixed-methods approach. Students' scores of the essays written in the pre-test and post-test phases constituted the quantitative part of the study, and the data elicited through semi-structured interviews with Iranian high school EFL students regarding the impact of LALE on their attitudes toward AaL formed the qualitative part.

3.2. Participants and Setting

The present study participants were selected from 120 male and female high-school EFL learners in Hejrat High School in Najafabad whose ages ranged between 16 and 18 (M=16.5, SD=.93). Oxford Placement Test (OPT) was administered to homogenize students. Based on their performance on the OPT, 80 participants with scores placed one standard deviation above and below the mean were chosen as the main intermediate-level participants of the study. They were then randomly categorized into experimental and control groups, each with 40 students. The participants were requested to sign a consent form to take part in the study.

3.3. Instrumentation

3.3.1 Oxford Placement Test (OPT). The Oxford Placement Test, developed by Allan (1985), was used in the current study to homogenize the participants. This test is considered a global, efficient, and reliable means of placing EFL learners at different language proficiency levels. It is quick and easy to administer and takes approximately 60 minutes. Those students whose scores fell between 51-59 (as cut-off scores) were selected as intermediate-level students (Allan, 1985).

3.3.2 The Writing Scale (ESL Composition Profile). the ESL Composition Scale (Jacobs et al., 1981) was used as a reliable writing scale to teach the experimental group self- and peer-assessment techniques and strategies. It should be noted that the writing scale was not used in its original form because it just included the scoring rubrics and brief descriptions of the keywords and writing components (grammar, vocabulary, content, organization, and mechanics). Therefore, the instructor developed a pamphlet in Persian, providing detailed, less technical, and simple explanations of the keywords and writing components mentioned in the writing scale, accompanied by some examples. Moreover, the instructor provided these explanations to the participants orally and in simple language.

3.3.3. Interview on Attitudes Toward Assessment Practices. it was a researcher-made semi-structured interview that included six questions on learners' attitudes (Appendix A). The interview aimed to explore the positive and negative attitudes toward assessment practices as learning based on the Western and Northern Canadian Protocol for Collaboration in Basic Education (WNCP, 2006) model. All interview sessions were recorded and transcribed for content analysis.

3.4. Procedures

From 120 male and female high-school EFL learners, 80 homogenized students at the intermediate proficiency level were chosen based on the results of the OPT test and were then randomly categorized into experimental and control groups, each with 40 students. To gain evaluative, critical judgment on the quality of their own or their peers' work, to compare their works with those of peers of the same status, to understand the criteria and standards of scoring, to assess themselves based on standard rating scales, and to have an active role in learning through assessment and reflect upon what they can do; first, both groups were requested to write an essay on the topic of: "Neighbors are those who live near us. What qualities a good neighbor should have? Specific details and examples are required." and assess the writings of other students and give them feedback in pre-test phase. They were supposed to write an essay of around 500 words in about 90 minutes. The students were permitted to use their dictionaries to write essays.

Afterward, the ESL Composition Scale (Jacobs et al., 1981) (Appendix B) was used as the writing scale to teach the experimental group the self- and peer-assessment techniques and strategies. In this study, the writing rubric was used both as an assessment and an instructional tool. It is worth mentioning that the writing scale was not used in its original form because it just involved the scoring

rubrics and brief descriptions of the keywords and writing components (grammar, vocabulary, content, organization, and mechanics). Therefore, the instructor developed a pamphlet in Persian including detailed, less technical, and simple explanations of the keywords and writing components mentioned in the writing scale, accompanied by some examples to let the students know what makes good writing and what criteria and standards they are supposed to learn to be able to write a good essay. Considering the proficiency level of the participants, these pamphlets (Appendix B) were also given to ensure that everything was utterly comprehensible to them. Moreover, the instructor provided these same explanations to the participants orally and in simple language during 13 treatment sessions.

In addition, pamphlets given to the participants included sample paragraphs from students who had previously taken a writing course, and raters rated their paragraphs based on the writing rubric. The aim was to give the students informative feedback and detailed evaluation of their peers' writing works to help them understand the rating scale more accurately and, consequently, learn through assessment and rating. During the 13 treatment sessions, the students were also asked to assess their own and their peers' writings according to the provided rubric to support student learning and the progress of sophisticated thinking skills. This kind of peer- and self-assessment based on the provided rubric is intended to develop the students learning and metacognition, i.e., self-monitoring and self-regulating (Goodrich, 1996).

The control group received no instructional information on assessment rubrics. In the post-test phase, one week after the treatment, both groups of students were asked to write an essay on another specific topic to find if the treatment (language assessment enhancement) has influenced the students' knowledge of assessment as learning of the writing skill. The topic was: Explain the various effects of television on people's behavior. You are supposed to provide specific reasons and examples.

In this rubric, a four-point scale was used for each of the five criteria. Each scale, devoting a column to itself (from A to D), specified a different performance level within those criteria. The descriptions in the first column were each worth 4 points, the second column 3 points, the third two points, and the fourth one point. Considering, for instance, the grammar criteria, an A paper lacks the most grammatical errors and would be worth four points. Paper B includes several errors but generally uses good grammar and is worth three points. Paper C involves frequent grammatical errors and is worth three points. Paper D is full of grammatical mistakes and is worth one point, showing that the student does not possess the appropriate grammatical knowledge for this level of language learning. Two raters were asked to read the essays twice to determine where on the scale the essay of each student fell for each of the criteria. The total scores of each student for each criterion of the rubric were calculated. Then, it was divided by the total criteria, five in this case, to decide upon the student's final grade. The result of the study revealed that there was internal consistency between the two raters when scoring the students' essays using the Jacobs ESL Composition Profile, and this internal consistency was high (r. = 0.67, $\alpha = 0.00 < 0.05$). The Cronbach alpha coefficient also was 0.72, indicating an appropriate level of reliability.

In the qualitative phase of the study, ten high school EFL students from the experimental group were interviewed about their attitudes toward the application of assessment as learning writing skill in their English classes. The researcher-made semi-structured interview included six questions on learners' attitudes toward assessment practices as learning based on the WNCP (2006). The first interview items were created based on an inclusive examination and evaluation of the related literature and the researcher's knowledge of language teaching and assessment in Iranian high schools. It is also worth mentioning that following the semi-structured interview protocol, the researcher designed and asked more open-ended questions to discuss with the interviewees, not just a straightforward question-and-answer format.

Considering the potential problems, the interview questions were piloted. Two EFL students agreed to participate in the interview piloting whose results helped the interviewer know how long the interview takes, the wording of questions (if they are ambiguous or confusing to the interviewees), and procedural matters such as the ordering and sequence of the questions and the design of the introduction. Moreover, to validate the interview questions, two associate professors of TEFL were asked to comment on them. The researcher made the required modifications to ensure this instrument could produce the rich data required for a valid thematic analysis.

Before starting the interview, the researcher called the selected participants to set a convenient time for the interview. At first, the researcher informed the participants about the purpose of the interview and its expected duration, and they were encouraged to give honest answers to the questions. The interviews were held face-to-face, and depending on the information provided, each participant was interviewed for 10 to 15 minutes in a quiet room in the high school. The interviews were held in Persian to ensure they understood the questions clearly. This would also help the participants feel more relaxed and freer while answering the questions and expressing their ideas better. An audio recorder was used to record the interview sessions.

All the recorded audio files were transcribed verbatim within a month for content analysis. Afterward, the researcher read the transcribed raw data several times to understand the participant's thoughts and ideas better and identify themes and categories in their responses. The researcher followed a coding system, using brief phrases as codes, renamed the initial codes, omitted the redundant ones, and merged those that denoted a similar concept. The codes merged into larger categories, and the lengthy interview transcript was turned into briefer and more concise formulations used to formulate assertions for each theme. Both frequent and non-frequent themes were interpreted based on the relevant literature or the researcher's understanding.

3.5. Data analysis

The quantitative data were analyzed by independent-samples t-test and repeated measures analysis of variance (ANOVA), and the interviews were analyzed using the content analysis method.

4. Results

Before starting the main analyses, it should be ensured that the obtained writing scores were normally distributed. Kolmogorov-Smirnov test was conducted to check the normality of writing scores, and the summary of results is demonstrated in Table 1 below.

Table 1
Kolmogorov-Smirnov Test of Normality for Speaking Pre-and Post-test

		Pre-test	Post-test
N		80	80
Normal Parameters	Mean	6.96	11.86
	SD	1.67	4.49
Most Extreme	Absolute	.15	.15
Differences	Positive	.15	.13
	Negative	-.12	-.15
Test Statistic		.15	.15
Asymp. Sig. (2-tailed)		.07	.09

The Kolmogorov-Smirnov test revealed that the participants' scores on writing pre-test ($W = .15$, $p = .07$) and post-test ($W = .15$, $p = .09$) were normally distributed (Table 1). Therefore, the distribution of the pre-test and post-test scores was normal.

The first research question sought to find whether language assessment literacy enhancement (LALE) had any significant effect on students' assessment as learning of writing skill in Iranian high school EFL classes compared with having no instruction. In so doing, two independent samples t-tests were run to compare the pre-and post-test scores of LALE and the control groups, the results of which are presented below.

Table 2
Descriptive Statistics of Pre-test Writing Scores of Both

	M	SD	N
LALE group	6.8	1.62	40
Control group	7.1	1.76	40

As shown in Table 2, the pre-test scores of the experimental and control groups are fairly close to each other. Yet, an independent samples t-test was run to compare the pre-test writing scores of both groups.

Table 3
Independent Samples T-Test for Pre-test Writing 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	.5	.47	1.84	78	.06	.7	.37	-.05	1.45
Equal variances not assumed			1.84	77.4	.06	.7	.37	-.05	1.45

There was no significant difference between the pre-test scores of LALE (M = 6.8, SD = 1.62) and control (M = 7.1, SD = 1.76; t (78) = 1.84, p = .06, two-tailed) (Table 3). It can be concluded that both groups were equal in terms of writing ability before the intervention.

Furthermore, an independent samples t-test was run to compare the post-test writing scores of the LALE and the control group.

Table 4
Descriptive Statistics of Post-Test Writing Scores of Both Groups

Group	M	SD	N
LALE group	15.77	1.8	40
Control group	7.35	1.8	40

The experimental (LALE) group obtained higher scores in the post-test of writing (Table 4). An independent samples t-test was run to see if the difference between the two groups was significant.

Table 5
Independent Samples T-Test for Writing 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	.19	.04	20.87	78	.00	8.42	.4	7.62	9.22
Equal variances not assumed			20.87	78	.00	8.42	.4	7.62	9.22

There was a significant difference between post-test writing scores of experimental (M = 15.77, SD = 1.8) and control (M = 7.35, SD = 1.8) groups; t (78) = 20.87, p = .00, two-tailed) (Table 5). The magnitude of the differences in the means was large (eta squared = .82). In general, the experimental (LALE) group outperformed the control group in the writing post-test.

The second research question aimed to find the writing skill (i.e., grammar, vocabulary, mechanics, organization, and content) that was more significantly influenced by language assessment literacy enhancement (LALE). A repeated measure analysis of variance ANOVA was run to answer this question.

Table 6
Descriptive Statistics of Pre- and Post-Test Writing Skills' Scores of LALE Group

	M	SD	N
Grammar pre-test	2.1	1.04	40
Vocabulary pre-test	1.58	.67	40
Mechanics pre-test	1.07	.26	40
Organization pre-test	1	.00	40
Content pre-test	1	.00	40
Grammar post-test	3.79	.4	40
Vocabulary post-test	3.58	.49	40
Mechanics post-test	3.05	.6	40
Organization post-test	2.82	.6	40
Content post-test	2.5	.5	40

The means of the pre-and post-test scores of writing skills were different from pre-to post-test (Table 6). However, the multivariate test result needs to be checked to find if the differences were statistically significant (Table 7).

Table 7
Multivariate Test for Writing Skills in LALE Group

Effect	Value	F	Hypothesis df	Error df	Sig.
Pillai's Trace	.98	286.36	8	31	.00
Wilks' Lambda	.01	286.36	8	31	.00
Hotelling's Trace	73.9	286.36	8	31	.00
Roy's Largest Root	73.9	286.36	8	31	.00

As shown in Table 7, the result of Wilk's Lambda F (8, 31) = 286.36, p= .00 reveals a statistically significant difference among the scores of writing skills. The pairwise comparison results (Table 8) show the components whose difference between pre-and post-test scores was significant.

Table 8
Pairwise Comparisons of Writing Skill Scores of LALE Group

Component	Mean difference	Std. Error	Sig.
Grammar (post- and pre-test)	1.62*	.17	.00
Vocabulary (post- and pre-test)	2*	.12	.00
Mechanics (post- and pre-test)	1.97*	.11	.00
Organization (post- and pre-test)	1.82*	.09	.00
Content (post- and pre-test)	1.51*	.08	.00

*mean difference is significant at the .05 level

The pairwise comparisons table indicates that the difference between the mean of pre-and post-test scores of writing skills was significant for the experimental group (p< .05). To clarify, language assessment literacy enhancement (LALE) instruction increased the writing skills' scores. Moreover, among the five writing skills, vocabulary (M= 2, SD= .12) was more significantly influenced by LALE instruction.

For the third research question, the content analysis of the interviews helped to extract three main themes shaping the participants' attitudes, namely 'lack of practice of self- and peer-assessment,' 'lack of confidence about the ability to use self- and peer-assessment,' and 'uncertainty about self- and peer-assessment as reliable tools for measurement.' Each of these themes is introduced below, and related excerpts from interview transcripts are provided as well.

The first theme: Lack of practice of self-and peer-assessment: a majority of the interviewees (90.2%) were not satisfied with the current amount of self-and peer-assessment as indicators of assessment as learning knowledge in their English classes. Most of them claimed that they had not received such kind of instruction in their English classes as the teacher is the main authority in the class, and the students are not usually allowed to actively participate in the evaluation and assessment process in the Iranian EFL context. As the analysis of the transcribed interviews revealed, almost all interviewees were unfamiliar with self- and peer-assessment techniques and had received no instruction in language assessment literacy. They had never been previously informed about the criteria for grading a written work. Below are two excerpts supporting this theme.

Student A: We were completely unfamiliar with self- and peer-assessment as we were never instructed or required to evaluate, assess, or score our own or peers' homework as part of the class activity. As you know, it is not usually a norm in Iran to let the students do the rating; it is considered the teacher's responsibility.

Student B: In the Iranian EFL classes, students are not allowed to do assessments and assign scores to their peers or classmates as the teacher is considered the only individual with the required knowledge and capability to assess the students and assign them scores. So, I can say that we rarely participate in self and peer-assessment activities.

The second theme: A lack of confidence in the ability to use self- and peer assessment: the majority of the students stated that, before being instructed, they believed that the teacher was the only person in the class who had the ability and deserved to provide feedback and give grades to students' writing. Two related excerpts are provided here:

Student A: I believe that my peers and I do not have the required knowledge and ability to assess, evaluate, score, or rate our own or our peers' writing, as it is a very difficult task that needs a lot of specialty and knowledge that we do not have.

Student B: I have neither the tendency and confidence nor the capability to rate and score my or my peers' work. I believe it is just the teacher who can accurately assess and score the students' writing.

The third theme: Uncertainty about self- and peer-assessment as reliable tools for measurement: another concern of the students before the instruction was their uncertainty about using self -and peer-assessment as reliable tools for measurement. They were worried if the scores they gave to their peers or their peers considered for them were reliable. Moreover, students stated that they felt uncomfortable to critically evaluate and rate the writing works of their peers as their identities were displayed. They did not like to criticize their peers' works in public. Therefore, there is the possibility that the students over-mark or under-mark their peers and provide them with unreliable feedback. Some of them even thought that the feedback provided by their peers was unfair. The following excerpt reflects this:

Student A: I do not like to talk critically about my friends' mistakes, especially the ones I am close to, in class. I think this makes them feel discomfort or shame. So, I have to ignore some of their mistakes and give them scores that they do not deserve. I also do not like my classmates being informed of my grades or my mistakes, specifically when I have made so many mistakes in my exams or homework. In such cases, my classmates may also give me scores higher than my real scores. That is why I think self- and peer-assessment are not good assessment tools.

5. Discussion

The present study aimed to investigate if language assessment literacy enhancement could significantly affect Iranian high school EFL learners' knowledge of assessment as learning. Several research studies, both in Iran and in European contexts, have been conducted to evaluate teachers' language assessment literacy from various aspects, especially considering demographic features such as age, gender, and years of teaching experience. However, studies examining the role of language assessment literacy enhancement (LALE) in developing the EFL learners' knowledge of assessment as learning is sporadic, if there is any. The innovative aspect of this study was examining the effect of LALE on the EFL students' writing skill and their perception of LALE training.

The study results revealed a significant improvement in students' knowledge of assessment as learning, represented in their essay writing skills, due to language assessment literacy enhancement. Black and Wiliam (2018) state that assessment as learning denotes the situation where learning and assessment can be integrated, and assessment mainly supports learning. LALE helped the participants become familiar with the writing assessment criteria, provide and receive feedback to and from their peers, think and reflect on their learning, and diagnose their strengths and weaknesses. The obtained results, in general, suggest that language assessment literacy enhancement could help Iranian EFL learners acquire a better awareness of the writing evaluation criteria and, consequently, they become more precise and accurate in assessing their own and their peers' writing ability, leading to the development of their writing skill.

Among the five writing skill criteria (i.e., grammar, vocabulary, mechanics, organization, and content), vocabulary was more significantly improved than the other criteria due to language assessment literacy enhancement. Vocabulary learning and memorization of word lists have been one of the main objectives of Iranian EFL learners and teachers for many years in Iran's educational context, where teaching discrete bits of language knowledge, especially vocabulary and grammar, has been the focus of attention for many years. Although the reformed curriculum in Iran focuses on speaking and listening, the instruction in the classroom and assessment practices have not significantly changed from the previous grammar translation method. That is probably why students are familiar with and perform better in vocabulary use compared to other writing criteria.

Having generated a significant impact hinges on some factors, such as the student's attitude and their perception of self- and peer-assessment. Student attitude is considered important in successfully implementing peer assessment as it specifies the quality of self- and peer-assessment by the feedback level and the accurate scoring. Therefore, the study aimed to investigate if language assessment literacy enhancement (LALE) affects students' attitudes toward assessment as learning in Iranian high school EFL classes. As the study results indicated, Iranian high school EFL students commonly hold positive attitudes toward the impact of peer-and self-assessment on developing their writing skills. However, they had some negative attitudes regarding self and peer-assessment implementation in the Iranian high school EFL classes. The content analysis of the interviews helped to extract three main themes shaping the participants' attitudes, namely 'lack of practice of self-and peer-assessment,' 'lack of confidence about the ability to use self- and peer-assessment,' and 'uncertainty about self- and peer-assessment as reliable tools for measurement.'

Regarding the first theme of the study, lack of practice of self- and peer-assessment, as stated by Adiyani (2021), previous studies revealed that the addition of frequency of use and taking part in peer-assessment activities leads to the promotion of positive change in the attitude of students. Including peer-assessment in writing classrooms will help students develop a positive attitude and give them the required and essential skills that primarily restrained their motivation to take part in the writing peer assessment. Moreover, the implementation of peer assessment in the writing classrooms will make the students develop their high-order writing and thinking skills since they can provide an objective judgment of their peers' works and constantly revise, modify, and contemplate their works based on the assessment provided by their peers. Ashenafi (2015) believes frequent peer assessment activities can decrease students' negative attitudes. This signifies that in classes where peer assessment is used as an integral and basic part of their assessment, a student's negative attitude would become positive after frequent participation. Previous research studies suggested that the inclusion of peer assessment promotes positive change in a student's attitude (Adiyani, 2021).

Concerning the second theme of the study, lack of confidence about the ability to use self- and peer-assessment, students did not have confidence in their ability to play the role of assessor, which aligns with the previous studies (e.g., Birjandi & Bolghari, 2015; Landry et al., 2014; Zou et al., 2017) pointing out that EFL students doubt their own and their peers' competences in giving helpful feedback and an unbiased score. In addition, the students who previously underwent no responsibility for their own or their peers' assessments are now supposed to be responsible for their learning through self-assessment and self-regulation (Razavipour & Rezagah, 2018). To explain it, students thought they did not have the required experience and knowledge to properly grade/ rate their peers' work, leading to a negative attitude toward self and peer assessment.

Regarding the third theme, uncertainty about self- and peer-assessment as reliable tools for measurement and that students felt uncomfortable critically evaluating and rating the writing works of their peers as their identities were displayed, one possible solution is to be anonymous. Regarding peer assessment, Wang et al. (2019) showed that university students want to develop and have a positive attitude toward peer-assessment, provided that the process lets them assess and evaluate their peers anonymously. Students' belief that the possibility of reliable, fair, and honest reviews is higher when identities are not known reflects their positive attitude toward anonymous peer assessment. This is supported by the findings of Zou et al. (2017), who showed that nearly half of the subjects have prior familiarity with peer-assessment by online anonymous peer review. Lin (2018) noted that a lack of familiarity with criticism could also cause the students to show constructive feedback, mainly negative rather than positive.

6. Conclusion

The study findings indicated significant improvement in students' knowledge of assessment as learning, represented in their essay writing skills, as the result of language assessment literacy enhancement (LALE) as it involved the participants in giving and receiving feedback to and from their peers, in thinking and reflecting on their learning, and in diagnosing their strengths and weaknesses. The content analysis of reviews with the EFL students regarding their attitudes about the language assessment literacy enhancement and its impact on their knowledge of assessment resulted in three main themes: uncertainty, lack of practice, and unreliable rating. They were dissatisfied with the amount and frequency of language assessment instruction in their EFL classes, the lack of familiarity with the grading rubrics, and the lack of knowledge of assessment as learning techniques and strategies. They did not feel comfortable to explicitly and critically evaluate their peers. They also felt ashamed when their work was critically graded by their classmates. One possible solution to this can be anonymous evaluation and rating. Moreover, the Iranian EFL students did not have enough confidence in their ability to be the assessors of themselves or their peers. They had relied on their teachers as the only reliable assessors for many years and were uncertain about their ability to judge, evaluate, and rate their peers accurately.

This study just focused on the writing skill of the learners due to the limited time and scope of the study; however, other skills and language components can be investigated as well to examine the impact of LAL on their development through assessment as learning. Moreover, the impact of developing and strengthening assessment as learning knowledge on EFL learners at other proficiency levels, especially on more proficient learners, can be a good subject of study.

The present study findings can be useful both for EFL learners and teachers. EFL learners can benefit from self- and peer-assessment strategies and techniques for better learning. EFL teachers can use them to create more learner-centered activities and help their students become independent learners, providing them with the required criteria and instructions.

Declaration of Conflicting Interests

We declare that this manuscript is original, has not been published before, and is not currently being considered for publication elsewhere. We know of no conflicts of interest associated with this publication.

Funding

We declare that there has been no financial support for this work.

References

Adiyani, P. N. (2021). EFL students' attitude towards writing peer assessment: A systematic review. *Research on English Language Teaching in Indonesia, 9*(3), 1-10. https://ejournal.unesa.ac.id/index.php/retain/article/view/38432

Allan, D. (1985). Distinctions and dichotomy article for ETP testing and assessment. *Testing and Assessment in ELT*, *41*(3), 19-31.

Ashenafi, M. M. (2015). Peer-assessment in higher education twenty-first-century practices, challenges, and the way forward. *Assessment and Evaluation in Higher Education, 42*(2), 1-26. https://doi.org/10.1080/02602938.2015.1100711

Aslanoglu (2022). Examining the effect of peer and self-assessment practices on writing skills. *International Journal of Assessment Tools in Education*, *9,*179-196. https://doi.org/10.21449/ijate.1127815

Birjandi, P., & Bolghari, M. S. (2015). The relationship between the accuracy of self-and peer-assessment of Iranian intermediate EFL learners and their learning styles. *Theory & Practice in Language Studies*, *5*(5), 1–18. http://dx.doi.org/10.17507/tpls.0505.15

Black, P. & William, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice, 25*(6)*,* 551-575. https://doi.org/10.1080/0969594X.2018.1441807

Chen, W. (2023). *Improving the effectiveness of assessment as learning through an intervention*. (Doctoral dissertation), Hong Kong Polytechnic University.

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

Crooks, T. (2001). *The validity of formative assessment*. Paper presented to the British Educational Research Association Annual Conference, University of Leeds.

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

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

Effendi, T. (2020). *Language assessment: Exploring some issues in language assessment*. UNJ Press.

Fahimi, Z., & Rahimi, A. (2015). On the impact of self-assessment practice on writing skill. *Procedia: Social and Behavioral Sciences, 192,* 730–736. http://doi.org/10.1016/j.sbspro.2015.06.082

Fathi, J, Afzali, M, & Parsa, K. (2021). Self-assessment and peer-assessment in EFL context: An investigation of writing performance and writing self-efficacy. *Critical Literary Studies, 3*(1), 211-232. https://sid.ir/paper/782195/en

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.

Goodrich, H. (1996). *Student self-assessment: At the intersection of metacognition and authentic assessment*. (Doctoral dissertation), Harvard University.

He, M., & Wang, L. (2024). Implementing assessment as learning in online EFL writing classes. *RELC Journal,* 00336882231224966. https://doi.org/10.1177/00336882231224966

Hedayati, N., & Khoorsand, M. (2024). Tapping into assessment literacy: Unfolding assessment as learning strategies and EFL learners' writing development. *REiLA: Journal of Research and Innovation in Language, 6*(1), 21-37. https://doi.org/10.31849/reila.v6i1.15894

Jacobs, H. L., S. A. Zingraf, D. R. Wormuth, V. F. Hartfiel, & J. B. Hughey. (1981). *Testing ESL composition: A practical approach.* Newbury House.

Landry, A., Jacobs, S., & Newton, G. (2014). Effective use of peer assessment in a graduate level writing assignment: A case study. *International Journal of Higher Education, 4*(1), 38-51. https://doi.org/10.5430/ijhe.v4n1p38

Lin, G. Y. (2018). Anonymous versus identified peer assessment via a Facebook-based learning application: Effects on quality of peer feedback, perceived learning, perceived fairness, and attitude toward the system. *Computers & Education, 116*, 81-92. https://doi.org/10.1016/j.compedu.2017.08.010

Mohammadi, R., Ghanbari, N., & Abbasi, A. (2024). Perceptual (mis)matches between learners' and teachers' rating criteria in the Iranian EFL writing self-assessment context. *International Journal of Language Testing, 14*(1), 150–165. https://doi.org/10.22034/IJLT.2023.410811.1274

Naeini, J. (2013). *Graduated prompts and mediated learning experience: A comparative study of the effects of two approaches of dynamic assessment on the reading comprehension of Iranian EFL learners* (Doctoral dissertation). Islamic Azad University, Science and Research Branch, Tehran, Iran.

Ndoye, A. (2017). Peer /Self-assessment and student learning. *International Journal of Teaching and Learning in Higher Education, 29*(2), 255-269. http://www.isetl.org/ijtlhe/ ISSN 1812-9129

Nemati, M. & Ghafoori, M. (2016). The comparative effect of self-assessment, peer assessment and teacher assessment on the EFL learners' second language writing Development*. Linguistic Research in Foreign Languages*, *6*(1), 113-93. https://doi.org/10.22059/JFLR.2016.62811

Razavipour, K., & Rezagah, K. (2018). Language assessment in the new English curriculum in Iran: Managerial, institutional, and professional barriers. *Language Testing in Asia, 8*(1), 1–18. https://doi.org/10.1186/s40468-018-0061-8

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119-144. https://doi.org/10.1007/BF00117714

Vygotsky, L. (1978). *Mind in society*. Harvard University Press.

Wang, L., Lee, I., & Park, M. (2020). Chinese university EFL teachers' beliefs and practices of classroom writing assessment. *Studies in Educational Evaluation*, *66*, 100890. https://doi.org/10.1016/j.stueduc.2020.100890

Western and Northern Canadian Protocol (WNCP) (2006). *Rethinking classroom assessment with purpose in mind: Assessment for learning, assessment as learning, assessment of learning*. Manitoba Education, Citizenship and Youth.

Zou,Y., Schun, Ch, D., Wang, Y., & Zhang, F. (2017). Student attitudes that predict participation in peer assessment. *Assessment and Evaluation in Higher Education, 43*(5), 800-811. https://doi.org/10.1080/02602938.2017.140987

Appendix A
Interview on Learners' Attitudes toward Assessment Practices
(Researcher-made)
1. What is your attitude about using self-assessment practices and strategies in your EFL writing courses?
2. What is your attitude about using peer-assessment practices and strategies by your English teachers in your EFL writing courses?
3. Do your English teachers use self- and peer-assessment tasks and practices in your EFL classes? If yes, how frequently?
4. In your opinion, what are the advantages and disadvantages of self- and peer-assessment practices?
5. Have you ever experienced self- and peer-assessment strategies and practices in writing skill classes?
6. What is your idea about Jacobs' ESL composition scale, its constituent criteria, and its benefits for the development of ESL writing skill?

Appendix B
the ESL Composition Scale (Jacobs et al., 1981)
Essay writing rubric

criteria	A Worth of 4 points	B Worth of 3 points	C Worth of 2 points	D Worth of 1 point
Grammar (tense, number, agreement, word order, function, articles, pronouns, prepositions)	Free of most grammatical errors	Some grammatical mistakes but generally showing successful grammar usage	Frequent grammatical errors	Dominated by grammatical mistakes
Vocabulary (knowledge of word, idiom, word form, choice, usage)	Mastery of vocabulary knowledge	Occasional errors of word, idiom, word form, choice, and usage	Frequent errors of word, idiom, word form, choice, and usage	Little knowledge of English word, idiom, word form, choice, and usage
Mechanics (spelling, punctuation, capitalization, paragraphing, hand-writing)	Few numbers of errors	Occasional errors	Frequent errors	Dominated by errors of mechanics
Organization	Clear expression of ideas, well-organized, logical sequence, cohesion	Loosely organized, but main ideas stand out logical but incomplete sequencing	Ideas confused or disconnected, lack of logical sequencing and development	No organization, not enough to evaluate
Content	Thorough development of ideas, completely relevant to the assigned topic	Limited development of ideas, mostly relevant to the assigned topic	inadequate development of ideas, weakly relevant to the assigned topic	Not enough to evaluate