International Journal of Research in English Education (IJREE)

Original Article **Published online: 20 March 2021.**

The Effect of Teacher vs. Learner-Assessment Activities on the Iranian Intermediate EFL Learners' Writing Ability

Neda Movahedi[1] & Ghasem Aghajanzadeh Kiasi[1*]

* Correspondence:
aghajanzadeh1970@yahoo.com
1. Department of English Language, College of Humanities, Rasht Branch, Islamic Azad University, Rasht, Iran

Received: 25 October 2020
Revision: 1 January 2021
Accepted: 19 January 2021
Published online: 20 March 2021

Abstract

The present study sought to investigate the effectiveness of teacher versus learner assessment activities on the Iranian intermediate EFL learners' writing ability. In order to find an answer to the questions of this research, 30 intermediate students from Zaban Iran Language Institute in Rasht, Iran were selected via administering Solutions Placement Test (SPT). Next, they were divided into three groups of 10. Learners were randomly assigned into two experimental groups of peer assessment and self-assessment, and one control group, teacher assessment. A pre-test of writing was then administered before the groups received 8 sessions of treatment through peer, self-, and teacher assessment techniques. After the treatment period, a post-test of writing was administered to all groups. The results of descriptive and inferential analyses revealed that the peer assessment group attained the highest scores on the writing test, and a statistically significant difference among the effects of the teacher assessment, peer assessment, and self-assessment on Iranian Intermediate EFL students' writing ability was reported. Also, the group using teacher assessment technique attained higher scores than the group that used the self-assessment technique. The implications of this study can be considered for EFL teachers, learners, materials developers, and syllabus designers.

Keywords: peer assessment, self-assessment, teacher assessment, writing

[Downloaded from ijreeonline.com on 2026-09-06]

[DOR: 20.1001.1.25384015.2021.6.1.3.5]
[DOI: 10.52547/ijree.6.1.49]

[Downloaded from ijreeonline.com on 2026-09-06]

[DOR: 20.1001.1.25384015.2021.6.1.3.5]

[DOI: 10.52547/ijree.6.1.49]

1. Introduction

Assessment has been defined variously in the literature. Among the many, Linn and Miller (2005) define assessment of student learning as a systematic process of collecting information about student progress towards the learning goals. They maintain that students' performance can be measured in various ways, including "traditional paper and pencil tests, extended responses (essays), performance of authentic task, teacher observation, and student self-report" (p. 75). Similarly, Dhindsa, Omar, and Waldrip (2007) characterize assessment as a key component of teaching and learning, "a systematic process of data gathering" about students' progress (p. 1261).

Learners' assessment influences how they learn, and teachers play a powerful role in those perceptions (Schut, Driessen, van Tartwijk, van der Vleuten, & Heeneman, 2018; Watling & Ginsburg, 2019; Zare Toofan, Vaseghi, & Zare, 2019). Teaching happens through human interaction, and, therefore, the characteristics of teachers' interaction and relationships with learners can make a substantial difference to the kind of learning environment they create (Ramani, Konings, Mann, Pisarski, & van der Vleuten, 2018; Telio, Ajjawi, & Regehr, 2015).

Assessment can also be classified into three types according to the person who evaluates. The three types of assessment are self-, peer, and teacher assessment. Self-assessment, according to Dikel (2009), refers to the judgment and assessment made by the students themselves for their own learning. Self-assessment, as one type of alternative assessment, with the increased attention to learner-centered curricula, needs analysis, and learner autonomy has gained popularity in recent years and its potential value as an instructional tool to facilitate learning as well as a measurement tool has been a topic of much discussion (Butler & Lee, 2010). The increasing interest in such methods has been driven by increasing recognition of the need to engage learners as active participants in the learning process and to equip them with the skills required to be effective life-long learners (Wood, 2009).

In peer assessment, however, the students are involved in rating and assessing each other's works. It provides the opportunity for the students to develop the responsibility and also the ability of judgment throughout the process of exchanging the fair and accurate feedbacks (Brown & Glasner, 2007) and having comparisons (Liu & Carless, 2006) with one another which are connected to the final result. Peer assessment, as a source and tool that facilitates the cooperative learning, allows the students to be involved in assessing each other's learning outputs and learning tasks, and in helping and scaffolding each other (Keith, 2005). They provide feedbacks to their peers according to their ability of judgment (Brown & Glasner, 2007) and there are benefits of it which is the promotion of the student learning through the cooperative learning.

In teacher assessment, the teachers are the ones who play an important role in students' learning by assessing their learning with the end in mind (Wiggins & McTighe, 2005). In Matsuno's (2009) view, teacher assessment refers to the assessment where the teacher is the one who assesses and evaluates the students' learning. The teacher monitors the students in terms of their process, performances, and the learning outcomes, and gives them feedback to improve their weaknesses. He draws our attention towards teachers' temporal traits and assessment standards in order to propose a substitute to psychometric assessment.

Similarly, according to Jones (2010), students should have a clear-cut outline of the goals they are moving towards a thorough understanding of the criteria they are assessed against and should be engaged in self- and peer assessment to enrich their work. It is necessary to make use of various types of alternative assessments, so as not only to assess learners' findings but also to promote the L2 teaching and learning.

Along with new developments in language teaching and assessment which try to enhance students' autonomy and learning, peer assessment and self-assessment are gaining momentum and are playing more significant role in language learning and teaching. They are procedures which give the students the right opportunity to judge their own learning because they need to know their own abilities, how much improvement they are making, and what they can do with the abilities they have achieved. As far as education is concerned, students' awareness of their own performance is really important. If they can do this appropriately enough, they may not be obliged to depend on the teachers' opinion.

Generally, assessment makes the students evaluate their learning throughout their learning process with an active involvement in learning and teaching process. In other words, it promotes reflective learning. It develops the students' evaluative skills and also provides an opportunity for them to be engaged to the reflective learning concurrently. According to Hinett (2012), the reflection that the students make for themselves in their own learning can improve

their understanding of learning and expand the quality and depth of learning by making them to judge and think of their own learning in the process of reflective learning.

1.1 Statement of the Problem

Assessment as an approach to teaching and learning that creates feedback is used to improve students' performance. Traditionally, teachers play a pivotal role in procedure of grading. However, they never apply the potential capability of students, and they are not involved, and cannot have any diagnostic feedback of what they have done on the test. In support of this, Weigle (as cited in Khodashenas & Rakhshi, 2017) believes that traditional approaches to writing assessment are incomplete because they are not able to assess learners' writing ability based on one draft which is written under strict timed conditions about an unfamiliar topic. On the other hand, the teachers are not in a position to make appropriate judgments about their students' writing assignment. This concept is in line with Nunan (2003) who believes that writing should be taught as a process rather than a product.

According to Pourverdi Vangah, Jafarpour, and Mohammadi (2016), in traditional methods of assessing writing, the teachers act like a reader and an editor; first they read the paper and then edit it for grammatical and mechanical mistakes while it contrasts with Brown and Harris (2013) who believe in incorporating both formal and informal assessment techniques for monitoring learners' progress in writing. In traditional classroom, there is no collaborative assessment in classroom. It has diminutive role during assessing process. Student assessment as a part of collaborative assessment refers to the involvement of students and teacher during assessing process. This assessment is done in conference session between the students and teacher.

Student-centered learning becomes a pioneer of development of learning approach. In this approach, students' activities are important indicators in learning process and quality of learning product (Zohrabi, Torabi, & Baybourdiani, 2012). In the teaching and learning English, this approach links with flexible learning, experiential learning, and self-directed learning (Acat & Dönmez, 2009). According to these authors, in teacher-centered learning, teachers usually use particular textbooks, which are mostly grammar oriented and to compare the language structures of native and target languages. In this situation students tend to be more competitive and individualistic because they have less opportunity to think aloud or interact. The main problem is that testing is more common than using assessment in classes. We often imagine testing to be what happens at the end of the year but the assessments may vary in terms of the scores they provide, their timing, and their relevance. Students need to have the opportunity to demonstrate in rating their papers and understand what their real scores are.

1.2 Research Question

In light of what were presented above, the current study investigated the effectiveness of teacher vs. learner-assessment activities on the Iranian intermediate EFL learners' writing ability. More precisely, the study strived to answer the following question:

Is there any statistically significant difference between the effects of peer assessment, self-assessment, and teacher assessment activities on Iranian intermediate EFL learners' writing ability?

1.3 Research Hypothesis

According the research questioned mentioned above, the following research hypothesis was proposed for the current study:

There is not any statistically significant difference between the effects of peer assessment, self-assessment, and teacher assessment activities on Iranian intermediate EFL learners' writing ability.

2. Literature Review

With the arrival of sociocultural theory on the scene in 1986, learning is conceptualized as an enterprise shaped and reshaped through social interaction on the premise that the human mind is always mediated by virtue of interaction with self or others (Lantolf, 2000). Put it another way, SCT rests on the premise that higher forms of thinking are formed and enhanced through interaction in a social context, and then they are transferred from the social to individual level (Vygotsky, 1978). Vygotsky's (1978) SCT of human learning describes learning as a social process and the origination of human intelligence in society or culture. The major theme of Vygotsky's theoretical framework is that social interaction plays a fundamental role in the development of cognition. The value of self-assessment as a learning activity has been prominent in the literature for over 30 years (Falchikov & Boud, 1989), and has been shown to

[Downloaded from ijreeonline.com on 2026-09-06]
[DOR: 20.1001.1.25384015.2021.6.1.3.5]
[DOI: 10.52547/ijree.6.1.49]

encourage deep approaches to learning (Ozogul & Sullivan, 2007) and the acquisition of essential skills for life-long learning (Thomas, Martin & Pleasants, 2011), to be positively related to student achievement, in both review (Brown & Harris, 2013) and empirical (Jay & Owen, 2016) studies.

Assessment is a major constituent of academic learning processes, and teaching and learning outcome focus heavily on it. The consequence of that focus is that assessment has the potential to drive student learning (Suri & Krishnan, 2019). For most students, according to James et al. (2002), "assessment requirements literally define the curriculum" (p. 7); therefore, it provides a means by which teaching academics can help guide student learning and engagement to maximize the learning potential it harbors. Assessment activities have the capacity to inspire students to engage with criteria and evaluate their own performance, thereby becoming self-regulated learners (Tai, Ajjawi, Boud, Dawson, & Panadero, 2018; Villarroel, Bloxham, Bruna, Bruna, & Herrera-Seda, 2018).

There is abundant evidence in the literature attesting that the use of self- and peer assessment has the capacity to enhance student learning (Boud, 2013; Falchikov, 2013; Yan & Brown, 2017). Neary et al. (2014, p. 9) point out that "engagement is created through active collaboration amongst and between students and academics." While conventional assessment allows others to make judgments about competence and success (Ajjawi, Tai, Nghia, Boud, Johnson, & Patrick, 2019) and "limits the potential of learner development through assessment (Spiller, 2012, p. 2), by combining self- and peer evaluation, allows the learner to become a part of the teaching-learning-assessment process. Rezaee, Rahimi, and Mehrabi (2019) found that peer-dynamic assessment group of students outperformed non-DA group of students in learning English language structures. Their findings suggest that the offered mediations sensitive to the participants' ZPD by their peers in a collaborative setting could lead to substantial changes in their grammatical learning, which was is in harmony with the theoretical foundations upon which DA is built.

The findings of Javaherbakhsh's (2010) investigation of the effect of Iranian advanced level learners' self-assessment on their performance in writing in English indicated that the self-assessment treatment administered to the experimental group had a significant effect on the learners' performance on the post-test of writing. In addition, the research conducted by Chang, Tseng, and Lou (2011) on the consistency and difference of teacher-, student self- and peer assessment among 72 senior high school students in the context of web-based portfolio assessment revealed that self- and teacher assessment were discovered to be consistent. However, consistency was not found between self- and peer assessment as well as peer and teacher assessment. In analyzing their consistency with the final examination, self- and teacher assessment demonstrated high consistency. They concluded that the teacher- and self-assessment outcomes reflect student achievements appropriately and hence had sufficient validity.

Regarding individual versus collaborative pair work, Storch (2005) studied the effectiveness of collaborative writing on L2 argumentative essays. Her study analyzed both the final product of their writing task (in terms of fluency, accuracy, and complexity) as well as the nature of interaction during the task. The results revealed that collaboration among team members leads to many opportunities for idea exchanging and peer feedback. Meihami and Razmjoo (2016) investigated the challenges and solutions of utilizing self- and peer-assessment in writing courses. By conducting a series of interviews and open-ended questions, Meihami and Razmjoo delved into the perspectives of 11 English Language Teaching (ELT) teachers and 56 ELT students to find out their challenges and probable solutions to applying self- and peer-assessment in the writing classes. The results of their study indicated that ELT students thought of self and peer's subjective judgment, learners' lack of assessment literacy, and instructional problems as the main challenges of self- and peer-assessment in writing classes. Moreover, ELT teachers believed that the challenges were the feasibility to do these types of assessment and ELT teachers' lack of assessment literacy. The common solution posited by both ELT teachers and students was development in teachers' assessment literacy.

In a study carried by Boumediene, Berrahal, and Harji (2016), the effect of the peer and self-assessment methods on writing ability of third year foreign languages learners enrolled at a secondary school in Algeria was investigated. The findings of the study exhibited a remarkable improvement in English writing performance of the peer group. Indeed, their study indicated a significant increase in the peer group's use of writing processes as a result of the peer assessment method. They concluded that the peer assessment model is an effective instructional strategy as well as an evaluation tool. Further, it promotes the learners' English writing performance by focusing efforts on writing products as well as writing processes.

Peer assessment, on the other hand, has received mixed reactions in the literature. While both the positive and negative aspects of peer assessment have been canvassed previously (Kearney, Perkins, & Kennedy-Clarke, 2016), the authors

[Downloaded from ijreeonline.com on 2026-09-06]
[DOR: 20.1001.1.25384015.2021.6.1.3.5]
[DOI: 10.52547/ijree.6.1.49]

believe that as part of a formative process, despite culminating in a summative mark, peer assessment can be quite useful, and that most of the negativity can be mitigated by practice and the moderating effect of the grading system. In their recent meta-analysis on peer assessment, Li, Xiong, Hunter, Guo, and Tywoniw (2019) found that peer assessment in general has a nontrivial positive effect on students' learning performance, which confirms previous literature on the benefits of peer assessment for student learning.

The results of a study carried out by Khorami Fard and Derakhshi (2019) application of DA as an alternative procedure to standard testing has a positive effect on both test performance and writing linguistic accuracy of English as a Foreign Language (EFL) learners. In a study involving university students in Iran, Esfandiari and Myford (2013) compared severity of self-, peer, and teacher assessments in foreign language writing. They found that on average teacher assessors rated more severely while self-assessors rated more leniently. Peer assessors turned to be half way between those two assessor groups. They attributed these differences to the influence of cultural more on students' abilities to self-assess and peer assess.

In a study carried out by Mosmery and Barzegar (2015), the effects of using peer-, self-, and teacher-assessment on nurturing students' accuracy in writing skill were analyzed and furthermore, the researchers intended to perform the study at three levels of task complexity (simple, medium, and complex), to find out whether being engaged in more complex tasks will help students improve accuracy in writing. The results indicated that although all three methods of assessment led to the students' progress, however, Self-Assessment method was the most effective method and Teacher and Peer were the second and third respectively. Furthermore, although a rise in the complexity level led to increment in results consistently, the increment in the second level was much higher than the third one which was contributable to less concentration in the third level.

In a similar study, but in a different culture, Matsuno (2009) compared self-, and peer assessments with teacher assessments in university writing classes. He found that in comparison with self-, and peer assessors, teacher assessors were neither lenient nor severe. Peer assessors produced fewer biased interactions compared with the self-, and teacher assessors. Based on the results of the study, here commended that, in some contexts, peer assessment can be used in writing classes. He concluded that "self-assessment was somewhat idiosyncratic and therefore of limited utility as a part of formal assessment" (p.75). Matsuno's finding mirrors the finding of the study by Esfandiari and Myford (2013) in that, in both studies, self-assessors were found to be significantly more lenient than peer, and teacher assessors.

3. Methodology

3.1 Design of the Study

The present study followed a quasi-experimental design in which the participants of this study were selected non-randomly from intact classes, and they were designated to three different groups of teacher assessment, self-assessment, and peer assessment non-randomly. Moreover, regarding the procedure, the present study followed a pretest-treatment-posttest design in which the participants took a pre-test at the onset of the study before they received the intervention programs of assessment techniques, and they took a post-test to know the effect of the programs on their writing ability.

3.2 Participants

The main population of the present study consisted of 45 Persian native speakers taking an English course at Zaban Iran Language Institute in Rasht, Iran. However, having been homogenized by SPT, 30 students were selected at the intermediate level. All participants were female students aged from 15 to 25. The participants were equally assigned as one control group (teacher assessment, N=10), and two experimental groups (peer assessment, N= 10), and (self-assessment, N=10).

3.3 Instruments and Materials

The first instrument used in the present study was SPT which was administrated to 45 students in order to select a homogeneous sample of intermediate level students. This test contains 50 multiple choice items which assess participants' knowledge of grammar and vocabulary from elementary to intermediate levels, a reading test with 10 graded comprehension questions, and an optional writing task that assesses students' ability to produce the language. The total score is 50. Those students who gain between 0-20, 21-30, and over 30 in grammar and vocabulary are regarded as elementary, pre-intermediate, and intermediate students, respectively. In reading, the scale is different. The total score is 10. The students who gain between 0-4, 5-7, and over 8 are considered as elementary, pre-

[Downloaded from ijreeonline.com on 2026-09-06]
[DOR: 20.1001.1.25384015.2021.6.1.3.5]
[DOI: 10.52547/ijree.6.1.49]

intermediate, and intermediate students, respectively. The second instrument consisted of writing topics for pre-test and post-test. The pre-test consisted of a topic entitled as "*Write a description of a person or a friend you know."* The topic of writing for the post-test was entitled as "*Are you a pessimist or an optimist person? Elaborate."* This change of topic was done to control for the probable testing effect. The forth instrument was pertinent to the subjective nature of the writing skill that required the researchers to use a scoring rubric based on which the writing ability of the students could be reliably scored. In so doing, a scoring rubric developed by Brown (2007) was used to score the students' writings in the pre-test and post-test (see appendix A).

3.4 Data Collection Procedure

The students took SPT for the purpose of being homogenized according to their English language proficiency level. Based on SPT scale, 30 students who scored (31+) in grammar and vocabulary and (8+) in reading and writing sections of the test were selected to participate in this study. Next, the students were randomly divided into three groups of a control group and two experimental groups with 10 students in each group. Later, all groups sat for the pre-test of writing. The purpose of this test was to assess the initial students' writing ability. Next, the groups received treatments based on peer assessment and self-assessment and teacher assessment strategies. In the self-assessment group, the students scored their own writing based on the rubric already given to them. In the teacher assessment group, the teachers scored students' papers based on the same rubric. In the peer assessment group, students scored each other's papers based on the same rubric already given to them. During the treatment they asked question and shared their ideas together.

After the treatment finished, the students took a post-test with the similar topic to the pre-test one in order for the researchers to know the effect of three types of assessments on the students' writing ability. The students' papers in the pre-test and post-test were scored by the researchers based on the same rubric. In order to measure the inter-rater reliability of writing scores, the Spearman's Rank Order Correlation (rho) was used to provide the agreement of the two raters. The scores were first analyzed descriptively, and the standard deviations and the means scores of both pre-test and post-test writing were calculated.

3.5 Data Analysis

To make sure of the scoring reliability, both pre-test and post-test writings were rated by the same two raters in this study. The inter-rater reliability of the writing scores (each student writing paper was scored by two raters in control group) was achieved via the Spearman Rank-order correlation. Regarding the results of the tests, in descriptive statistics, the means and standard deviations were calculated to find the difference among the results of three groups. With regard to the inferential statistics, the participants' post-test scores were analyzed by using a one-way ANOVA to find a statistically significance difference at the generally accepted alpha level (p-value) of 0.05. A post-hoc Scheffe test was also employed to find out where the group differences occurred.

4. Results

The inter-rater reliability estimates for the pre-test scores of writing in the three groups, as shown in Table 1, was found to be .950, .956, and .922 which were highly positive and showed significantly acceptable correlations.

Table 1. Inter-rater correlation for the pre-test scores of the groups

Correlations		
teacher assessment group	Spearman's Correlation	.950
	Sig. (2-tailed)	.000
	N	10
self-assessment group	Spearman's Correlation	.956
	Sig. (2-tailed)	.000
	N	10

[Downloaded from ijreeonline.com on 2026-09-06]

[DOR: 20.1001.1.25384015.2021.6.1.3.5]
[DOI: 10.52547/ijree.6.1.49]

peer assessment group	Spearman's Correlation	.922
	Sig. (2-tailed)	.000
	N	10

Table 2 represents descriptive statistics for writing test scores of the three groups on the pre-test before the onset of intervention program and applying the treatment.

Table 2. Descriptive statistics of the pre-test of writing scores

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Self-assess G.	10	16.6667	3.61498	.67518	11.5147	15.3148	11.00	17.00
Teacher assess G.	10	15.9000	3.92282	.75467	11.5874	14.2186	10.00	16.00
Peer assess G.	10	15.4667	3.44560	.63145	10.8758	14.8210	10.00	16.00
Total	30	16.0441	3.66113	.40095	11.3259	14.7848	10.33	16.33

All groups obtained almost the same values on the pre-test. That is, the mean scores are highly close to each other implying that the three groups were homogeneous regarding their writing skill. Although the self-assessment group had a slightly better performance, the difference was not significant.

According to Table 3, the measures of inter-rater reliability estimates for the post-test scores in the teacher assessment, self-assessment, and peer assessment were .962, .972, and .912 that were significant and showed a highly positive correlation between the raters.

Table 3. Inter-rater correlation for the post-test scores of the assessment groups

Correlations		
teacher assessment group	Spearman's Correlation	.962
	Sig. (2-tailed)	.000
	N	10
self-assessment group	Spearman's Correlation	.972
	Sig. (2-tailed)	.000
	N	10
peer assessment group	Spearman's Correlation	.912
	Sig. (2-tailed)	.000
	N	10

[Downloaded from ijreeonline.com on 2026-09-06]
[DOR: 20.1001.1.25384015.2021.6.1.3.5]
[DOI: 10.52547/ijree.6.1.49]

Table 4 demonstrates descriptive statistics for writing test scores of the groups on the post-test. The values obtained for the experimental groups showed a considerable difference in range, minimum, maximum, sum, and mean compared to the control group.

Table 4. Descriptive statistics of the post-test of writing scores

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Self-assess G.	10	18.4667	3.92550	1.01356	16.2928	20.6405	12.00	20.00
Teacher assess G.	10	22.1333	4.08598	1.05500	19.8706	24.3961	17.00	24.00
Peer assess G.	10	24.2000	2.90812	.75087	22.5895	25.8105	20.00	25.00
Total	30	21.6000	4.80719	.71661	20.8224	23.6157	12.00	23.00

For the writing test administered as the post-test, the mean scores for two experimental groups of self- and peer assessment, and the control or teacher assessment group were 18.46, 24.20, and 22.13, respectively. It can also be seen that the mean scores of the post-test writing were different from the pre-test scores demonstrating the different levels of writing ability. The participants' pre-test and post-test scores were also analyzed using a one-way ANOVA to find a statistically significance difference at the generally accepted alpha level (p-value) of 0.05. The Levene's test for homogeneity of variances was estimated for the pre-test writing scores to test whether the variance in scores was the same for each of the three groups.

Table 5. Results of Levene's test for the pre-test of l2 writing scores

Levene Statistic	df1	df2	Sig.
.736	2	272	.343

According to the result (Table 5), the significance value (Sig.) for Levene's test is greater than .05. It means that the Sig. value of .343 is greater than .05. So, the assumption of homogeneity of variance was met for the pretest writing. Reporting the one-way ANOVA, as shown in Table 6, the three groups were not statistically different in terms of writing ability at the beginning of the study before the immense of the treatment sessions.

Table 6. Results of One-Way ANOVA for the writing pre-test scores of groups

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	262.121	2	125.459	.214	.387
Within Groups	36123.012	27	7.598		
Total	62321.138	29			

F-tests are used to statistically assess the equality of means in the analysis of variance. The F (2, 27) statistics is equal to 0.214 and the probability value is 0.387. Since the F statistics is smaller than the probability value, it can be concluded that no statistically difference in the performance of the three different groups at the beginning of the study

[Downloaded from ijreeonline.com on 2026-09-06]
[DOR: 20.1001.1.25384015.2021.6.1.3.5]
[DOI: 10.52547/ijree.6.1.49]

was reported. Running one-way ANOVA to the post-test writing scores of the groups, the researchers ran the Levene's test for homogeneity of variances for the post-test scores as shown in Table 7.

Table 7. Results of Levene's test for the writing posttest scores of groups

Levene Statistic	df1	df2	Sig.
.731	2	27	.354

According to the significance value for Levene's test (.354) that is greater than .05, the assumption of homogeneity of variance is met as well in the groups' writing post-test scores.

Table 8. ANOVA results for the post-test scores of L2 writing

	Sum of Squares	Df	Mean	F	Sig.
Between Groups	448.933	2	224.467	5.602	.000
Within Groups	567.867	27	13.521		
Total	1016.800	29			

The results showed statistically significant differences between the groups at the $p = .000$. It means that the F statistics (2, 27 = 5.602) used to assess the equality of means is greater than the probability value of 0.000. Therefore, it can be concluded that there were differences among the three different groups after the intervention programs. Furthermore, the effect size was calculated by dividing the sum of squares for between-groups (448.933) by the total sum of squares (1016.800). The resulting eta squared value was .4, which in Cohen's (1988) terms would be considered an almost medium effect size. The ANOVA test reported that there was an overall difference between the groups, but it did not report specific groups differed. Because post-hoc tests are run to confirm where the differences occurred between groups, they should only be run when there is an overall statistically significant difference in group means. Accordingly, the post-hoc Scheffe test as a post hoc of multiple pairwise comparisons was used to confirm where the differences occurred between groups (Table 9).

Table 9. Post Hoc Scheffe test for the post-test of L2 writing

(I) Study Groups	(J) Study Groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Self-assess G.	Teacher assess G.	-3.6666*	1.25987	.032	-3.0739	-.2594
	Peer assess G.	-5.7333*	1.25987	.000	-3.1406	-2.3261
Teacher assess G.	Self-assess G.	+3.6666*	1.25987	.032	.2594	3.0739
	Peer assess G.	+2.0666*	1.25987	.016	-3.4739	-.3594
Peer assess G.	Teacher assess G.	-2.0666*	1.25987	.000	2.3261	3.1406
	Self-assess G.	+5.7333*	1.25987	.016	.3594	3.4739

A one-way between-groups analysis of variance was conducted to explore the impact of self-, peer, and teacher assessment on the Iranian intermediate EFL learners' writing ability. There was a statistically significant difference at the $p < .05$ level in L2 writing scores for the three groups: F (2, 27) = 5.6, $p = .000$. The effect size, calculated using

[Downloaded from ijreeonline.com on 2026-09-06]
[DOR: 20.1001.1.25384015.2021.6.1.3.5]
[DOI: 10.52547/ijree.6.1.49]

eta squared, was .04. Post-hoc comparisons using the post-hoc Scheffe test indicated that the mean score for peer assessment group (M = 24.2, SD = 2.98) was significantly different from self-assessment group (M = 18.46, SD = 3.92) and teacher assessment groups (M = 22.13, SD = 4.08).

5. Discussion

The present research aimed at investigating the impact of teacher and learner assessment (peer and self-assessment) on the writing ability of the Iranian intermediate EFL learners. According to the results of the statistical analysis, the participants' overall performance on writing was poor before the study; the overall low means on the pre-test of the three groups assumed that the participants were not good at writing skill in general. However, the participants' knowledge of L2 writing ability improved significantly over the course of the experiment and treatment. The three treatment conditions resulted in varying degrees of learning achievements; that is, the three groups did not improve equally across the pre-intervention and post-intervention programs in which the L2 writing ability of the participants was measured. The peer assessment and teacher assessment groups gained greater achievement, but the peer assessment group benefited the most from assessment techniques.

The results of current study are mostly supported by the literature in the field. In some cases, however, the findings are opposed by some studies. For example, the findings are in line with the findings of Cheng and Warren (2005) who displayed the advantages of peer assessment in English language programs for undergraduate engineering students attending a university in Hong Kong. The results of their studies yielded that the students scoring their peers' papers showed higher language proficiency compared to the teacher assessment. That is why Boumediene, Berrahal, and Harji (2016) believe, based on their study, that the peer assessment model is an effective instructional strategy as well as an evaluation tool. Further, it promotes the learners' English writing performance by focusing efforts on writing products as well as writing processes.

Another study that partly supports the findings of the present study and partly opposes it is done by Chang, Tseng, and Lou (2011). Investigating the consistency and difference of teacher-, student self- and peer assessment among 72 senior high school students in the context of web-based portfolio assessment, they found that the results of self- and teacher assessment were consistent. However, consistency was not found between self- and peer assessment as well as peer and teacher assessment. Peer assessment showed a low level of consistency as opposed to finding of the present study. They concluded that the teacher- and self-assessment outcomes reflect students' achievements appropriately.

The effect of learners' self-assessment on their performance in writing showed that the self-assessment treatment had a significant effect on the learners' performance on the post-test of writing (Javaherbakhsh 2010). This finding is opposed to the finding of the present study as in the present study where the students in the self-assessment group were the poorest in the writing skill compared to peer assessment and teacher assessment groups. Similarly, the results are not compatible with the findings of the study conducted by Mosmery and Barzegar (2015), who found that although the results indicated that although all three methods of assessment led to the students' progress, self-assessment method was the most effective method and teacher and peer were the second and third respectively.

The results of the present study supported the idea that students' writing scores would improve by writing practices through peer-assessment. Accordingly, peer-assessment can be used in EFL writing classes as a technique whereby the LOA is put into practice. They can be combined to improve the development of EFL students' writing ability. The present study might have both micro implications in the form of in-class teaching and macro implications in the form of curriculum planning and development, instructional design, and policymaking. This study shed light on the status of teaching English and the effect of peer-assessment on writing in an EFL context. It provided additional insights into better identifying existing challenges regarding assessment and in taking a more realistic perspective in regards to the ELT situation in Iran. The results of the study carry the following pedagogical implications:

The first implication is that LOA can be put into practice successfully through peer-assessment. Based on the findings of the current study, the integration of LOA and peer-assessment is effective both in theory and in practice as it had a significantly positive effect on EFL learners' writing ability. Moreover, peer-assessment can provide valuable information about the participants' writing process. Therefore, peer-assessment can perform the role of an alternative method whenever traditional ways of teaching writing do not bring the expected effect. The result of current study is also persuasive for relevant authorities to consider this new aspect of dealing with writing instruction and assessment.

To integrate and employ peer-assessment, learners and even teachers need specific guidelines and supports. They might need a longer time to implement this new technique appropriately. In other words, in order for peer-assessment

[Downloaded from ijreeonline.com on 2026-09-06]
[DOR: 20.1001.1.25384015.2021.6.1.3.5]
[DOI: 10.52547/ijree.6.1.49]

to perform a constructive role, teacher training is definitely a major factor. Teachers need to be trained in applying an appropriate procedure for implementing an integrated approach. The findings of this study recommend textbook writers and material designers to pay special attention to representation of writing in terms of processes where writing is broken down to different stages. The peer-assessment lets learners practice writing in a process-oriented system.

6. Conclusion

The current study used a quantitative method to study the effect of three techniques of teacher assessment, peer assessment, and self-assessment on the Iranian intermediate EFL learners' writing ability. The results of data analyses showed that one experimental group (peer assessment) and the control group (teacher assessment) improved significantly from the pre-test to the post-test. However, the peer assessment group outperformed the other two groups. It is concluded that the peer assessment is more efficient in teaching writing than the other methods. Based on the findings of the present study, it can be concluded that peer assessment enhances L2 learning, and it brings a shift in students' roles from passive learners to active participants and a change in using learning strategies as a result of engaging in assessment. Furthermore, as Topping (2009) argues, students in peer assessment are critically engaged with the materials assessed, they compare and contrast performance with their peers, and they identify errors in their own knowledge. In addition, students in their interaction or communication of feedback, may use similar and familiar language that will reduce negative feelings of being evaluated by their teacher as an authority (Liu et al., 2016). Therefore, feedback characteristics practiced in the peer assessment programs deemed important by current theories of peer assessment play a significant in students' learning achievements.

References

Acat, B., & Dönmez, İ. (2009). To compare student centered education and teacher centered education in primary science and technology lesson in terms of learning environments. *Procedia Social and Behavioral Sciences, 1*(1), 1805-1809. https://doi.org/10.1016/j.sbspro.2009.01.320

Ajjawi, R., Tai, J., Nghia, T., Boud, D., Johnson, L., & Patrick, C. (2019). Aligning assessment with the needs of work-integrated learning: The challenges of authentic assessment in a complex context. *Assessment & Evaluation in Higher Education, 45*(2), 304-316. https://doi.org/10.1080/02602938.2019.1639613

Boud, D. (2013). *Enhancing learning through self-assessment*. London: Routledge.

Boumediene, H., Berrahal, F. K., & Harji, M. B. (2016). The effectiveness of portfolio assessment on EFL students' writing performance: The case of third year secondary students in Algeria. *Academic Journal of Interdisciplinary Studies*, *5*(3), 119-127. https://doi.org/10.5901/ajis.2016.v5n3s1p119

Brown, S., & Glasner, A. (2007). *Evaluar en la Universidad. Problemas y nuevos enfoques.* Madrid. Narcea S. A. Ediciones.

Brown, G. T., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367-393). Sage, CA: Thousand Oaks.

Butler, Y. G., & Lee. J. (2010). The effects of self-assessment among young learners of English**.** *Language Testing, 27*(1), 5-31. doi: 10.1177/0265532209346370

Chang, C. C., Tseng, K. H., & Lou, S. J. (2011). A comparative analysis of the consistency and difference among teacher assessment, student self-assessment and peer assessment in a web-based portfolio assessment environment for high school students. *Computers and Education, 58*(1), 303-320. https://doi.org/10.1016/j.compedu.2011.08.005

Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing, 22*(1), 93-121. doi: 10.1191/0265532205lt298oa

Dhindsa, H., Omar, K., & Waldrip, B. (2007). Upper secondary Bruneian science students' perceptions of assessment. *International Journal of Science Education, 29*(10), 1261-1280. https://doi.org/10.1080/09500690600991149

Dikel, M. R. (2009). *A guide to going online for self-assessment tools*. https://www.vault.com/blogs/job-search/a-guide-to-going-online-for-self-assessment-tools

[Downloaded from ijreeonline.com on 2026-09-06]
[DOR: 20.1001.1.25384015.2021.6.1.3.5]
[DOI: 10.52547/ijree.6.1.49]

Esfandiari, R., & Myford, C. M. (2013). Severity differences among self-assessors, peer assessors, and teacher assessors rating EFL essays. *Assessing Writing, 18*(2), 111-131. https://doi.org/10.1016/j.asw.2012.12.002

Falchikov, N. (2013). *Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education*. London: Routledge.

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 59*(4), 395-430. https://doi.org/10.3102/00346543059004395

Hinett, K. (2002). *Improving learning through reflection-Part two*. https://www.researchgate.net/publication/228596017_Improving_learning_through_reflection-part_two

James, R., McInnis, C., & Devlin, M. (2002). *Assessing learning in Australian universities*. Melbourne: Centre for the Study of Higher Education, University of Melbourne.

Jay, J., & Owen, A. (2016). Providing opportunities for student self-assessment: The impact on the acquisition of psychomotor skills in occupational therapy students. *Assessment & Evaluation in Higher Education, 41*(8), 1176-1192. https://doi.org/10.1080/02602938.2015.1071317

Javaherbakhsh, M. R. (2010). The impact of self-assessment on Iranian EFL learners' writing skill. *English Language Teaching, 3*(2), 213- 218. https://doi.org/10.5539/elt.v3n2p213

Jones, J. (2010). The role of assessment for learning in the management of primary to secondary transition: Implications for language teachers. *Language Learning Journal, 38*(2), 175-191. https://doi.org/10.1080/09571730902928052

Kearney, S., Perkins, T., & Kennedy-Clark, S. (2016). Using self and peer-assessment for summative purposes: Analyzing the relative validity of the AASL (authentic assessment for sustainable learning) model. *Assessment and Evaluation in Higher Education, 41*(6), 840-853. https://doi.org/10.1080/02602938.2015.1039484

Keith, J. (2005). Trends in peer learning. *Educational Psychology, 25*(6), 631-645. https://doi.org/10.1080/01443410500345172

Khodashenas, M. R., & Rakhshi, F. (2017). The effect of electronic portfolio assessment on the writing performance of Iranian EFL learners. *International Journal of Research in English Education (IJREE), 2*(3), 67-71. https://doi.org/10.18869/acadpub.ijree.2.3.67

Khorami Fard, S., & Derakhshi, Z. (2019). On the role of dynamic assessment on promotion of writing linguistic accuracy among EFL learners: An interventionist model. *International Journal of Research in English Education (IJREE), 4*(2), 14-28. https://doi.org/10.29252/ijree.4.2.14

Lantolf, J. P. (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

Li, H., Xiong, Y., Hunter, C. V., Guo, X., & Tywoniw, R. (2019). Does peer assessment promote student learning? A meta-analysis. *Assessment & Evaluation in Higher Education, 45*(2), 193-211. https://doi.org/10.1080/02602938.2019.1620679

Linn, R. L., & Miller, M. D. (2005). *Measurement and assessment in teaching* (9thed.) Upper Saddle River, NJ: Prentice Hall.

Liu, N. F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education, 11*(3), 279-290. https://doi.org/10.1080/13562510600680582

Liu, C. C., Lu, K. H., Wu, L. Y., & Tsai, C. C. (2016). The impact of peer review on creative self-efficacy and learning performance in Web 2.0 learning activities. *Journal of Educational Technology & Society, 19*(2), 286-297. https://www.jstor.org/stable/jeductechsoci.19.2.286

Matsuno, S. (2009). Self-, peer, and teacher assessments in Japanese university EFL writing classrooms. *Language Testing, 26*(1), 75-100. https://doi.org/10.1177/0265532208097337

[Downloaded from ijreeonline.com on 2026-09-06]

[DOR: 20.1001.1.25384015.2021.6.1.3.5]

[DOI: 10.52547/ijree.6.1.49]

Meihami, H., & Razmjoo, S. A. (2016). An emic perspective toward challenges and solutions of self- and peer-assessment in writing courses. *Asian-Pacific Journal of Second and Foreign Language Education, 1*(9), 1-20. doi:10.1186/s40862-016-0014-7

Mosmery, P., & Barzegar, R. (2015). The effects of using peer, self and teacher-assessment on Iranian EFL learners' writing ability at three levels of task complexity. *International Journal of Research Studies in Language Learning, 4*(4), 15-27. https://doi.org/10.5861/ijrsll.2015.928

Neary, M., Saunders, G., Hagyard, A., & Derricott, D. (2014). *Student as producer: Research-engaged teaching: An institutional strategy*. New York: Higher Education Academy.

Nunan, D. (2003). *Second language teaching and learning*. Boston, Massachusetts: Heinle & Heinle Publishers.

Ozogul, G., & Sullivan, H. (2007). Student performance and attitudes under formative evaluation by teacher, self and peer evaluators. *Educational Technology Research and Development, 57*(3), 393-410. https://doi.org/10.1007/s11423-007-9052-7

Pourvandi Vangah, F., Jafarpur, M., & Mohammadi, M. (2016). Portfolio assessment and process writing: Its effect on EFL students' L2 writing. *Journal of Applied Linguistics and Language Research, 3*(3), 224-246. http://www.jallr.com/index.php/JALLR/article/view/307

Ramani, S., Konings, K. D., Mann, K. V., Pisarski, E. E., & van der Vleuten, C. P. M. (2018). About politeness, face, and feedback: Exploring resident and faculty perceptions of how institutional feedback culture influences feedback practices. *Academic Medicine, 93*(9), 1348-1358. https://doi.org/10.1097/ACM.0000000000002193

Rezaee, A., Rahimi, S., & Mehrabi, M. (2019). Cultivating grammar knowledge of EFL learners through informed peer-dynamic assessment. *International Journal of Research in English Education (IJREE), 4*(3), 70-83. https://doi.org/10.29252/ijree.4.3.70

Schut, S., Driessen, E., van Tartwijk, J., van der Vleuten, C., & Heeneman, S. (2018). Stakes in the eye of the beholder: An international study of learners' perceptions within programmatic assessment. *Medical Education, 52*(6), 654-663. https://doi.org/10.1111/medu.13532

Storch, N. (2005). Collaborative writing: Product, process and students' reflections. *Journal for Second Language Writing, 14*(3), 153-173. https://doi.org/10.1016/j.jslw.2005.05.002

Suri, H., & Krishnan, S. (2019). Assessment hurdles in core first year courses in Australian universities: Are we trying to catch out students? *Assessment & Evaluation in Higher Education, 45*(2), 251-265. https://doi.org/10.1080/02602938.2019.1632795

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgment: Enabling students to make decisions about the quality of work. *Higher Education, 76*(3), 467-481. https://doi.org/10.1007/s10734-017-0220-3

Telio, S., Ajjawi, R., & Regehr, G. (2015). The educational alliance as a framework for reconceptualizing feedback in medical education. *Journal of the Association of American Medical Colleges, 90*(5), 609-614. https://doi.org/10.1097/ACM.0000000000000560

Thomas, G., Martin, D., & Pleasant, K. (2011). Using self- and peer-assessment to enhance students' future-learning in higher education. *Journal of University Teaching & Learning Practice, 8*(1), 1-17. http://ro.uow.edu.au/jutlp/vol8/iss1/5

Topping, K. (2009). Peer assessment. *Theory into Practice, 48*(1), 20-27. https://doi.org/10.1080/00405840802577569

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education, 43*(5), 840-854. https://doi.org/10.1080/02602938.2017.1412396

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

[Downloaded from ijreeonline.com on 2026-09-06]

[DOR: 20.1001.1.25384015.2021.6.1.3.5]

[DOI: 10.52547/ijree.6.1.49]

Watling, C., & LaDonna, K. (2019). Where philosophy meets culture: Exploring how coaches conceptualize their roles. *Medical Education, 53*(5), 467-476. https://doi.org/10.1111/medu.13799

Wiggins, G., & McTighe, J. (2005). *Understanding by design: Expanded* (2nd ed.) Alexandria, VA: ASCD.

Wood, W. B. (2009). Innovations in teaching undergraduate biology and why we need them. *Annual Review of Cell and Developmental Biology, 25,* 93-112. https://doi.org/10.1146/annurev.cellbio.24.110707.175306

Yan, Z., & Brown, G. T. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment & Evaluation in Higher Education, 42*(8), 1247-1262. https://doi.org/10.1080/02602938.2016.1260091

Zare Toofan, Z., Vaseghi, R., & Zare, M. (2019). Iranian EFL learners' perceptions toward paper assessment in mid-term and final exams in an English language institute. *International Journal of Research in English Education (IJREE),* *4*(3), 21-41. http://ijreeonline.com/article-1-247-en.html

Zohrabi, M., Torabi, M. A., & Baybourdiani, P. (2012). Teacher-centered and/or student-centered learning: English language in Iran. *English Language and Literature Studies, 2*(3), 18-30. https://doi.org/10.5539/ells.v2n3p18

[Downloaded from ijreeonline.com on 2026-09-06]

[DOR: 20.1001.1.25384015.2021.6.1.3.5]

[DOI: 10.52547/ijree.6.1.49]

Appendix A

The Scoring Rubric Used to Score the Students' Writing Paper

Criteria	Descriptors	Scores
Focus	1. Having problems with focus or failing to address the writing task	1
	2. Inadequately addressing the writing task	2
	3. Addressing the writing task adequately but Sometimes straying from the task	3
	4. Addressing most of the writing task	4
	5. Specifically addressing the writing task	5
Elaboration	1. Using few or no details or irrelevant details to support topics or illustrate ideas	1
	2. Using inappropriate or insufficient details to support topics or illustrate ideas	2
	3. Using some details to support topics or illustrate ideas	3
	4. Using appropriate details to support topics or illustrate ideas	4
	5. Using specific appropriate details to support topics or illustrate ideas	5
Organization/ Support	1. The logical flow of ideas is not clear and connected	1
	2. The logical flow of ideas is less clear and connected	2
	3. The logical flow of ideas is mostly clear and connected	3
	4. The logical flow of ideas is generally clear and connected	4
	5. The logical flow of ideas is specifically clear and connected	5
Conventions (spelling, grammar and punctuation)	1. Standard English conventions are poor with frequent errors	1
	2. Standard English conventions are inappropriate with obvious errors	2
	3. Standard English conventions are fair with some minor errors	3
	4. Standard English conventions are almost accurate	4
	5. Standard English conventions are perfect or near perfect	5
Vocabulary	1. Little knowledge of English vocabulary, idioms and verb forms	1
	2. Frequent errors of word/idiom form, choice, usage; meaning confused or obscured	2
	3. Occasional errors of word/idiom form, choice, usage but meaning not Obscured	3
	4. Almost effective word/idiom form, choice, usage; almost appropriate register	4
	5. Effective word/idiom form, choice, usage; appropriate register	5

[Downloaded from ijreeonline.com on 2026-09-06]
[DOR: 20.1001.1.25384015.2021.6.1.3.5]
[DOI: 10.52547/ijree.6.1.49]