

The Effect of Peer Assessment on Improvement of Iranian Pre- Intermediate EFL Learners' Oral Production

Arezou Malek Gerdeh, Mehran Davaribina*

English language teaching department, Islamic Azad University Ardabil, Iran

Abstract

Many researchers are currently interested in employing new assessment technique such as peer and self-assessment. The fundamental goals of such strategies are to increase student's involvement in the classroom process and to share responsibility between teacher and students. The study's main aim was to discover the effect of applying peer assessment on the improvement of pre-intermediate Iranian EFL learner's oral production. The participants were randomly assigned into four groups (control and experimental) based on their gender, with 20 students in each group. The experimental groups received treatment in the form of peer assessment questionnaire while the control groups ran traditional sessions. Also the Oxford placement test and Cambridge Oral test (COT) was used as pre-and post-test of the study. The study followed true-experimental design. One-way ANOVA was run to measure the significant difference between proficiency levels before the main phase of the study. After finishing the data collection process, two-way ANOVA was applied to post-scores to see whether any improvement had taken place as a result of treatment. The findings indicated that there was statistically significant difference between experimental and control groups ($p = .000$), but there was not statistically significant difference between male and female in terms of the oral production gains ($p = .975$).

Keywords: Assessment, Oral production, Peer-assessment, Traditional-assessment, EFL

1. Introduction

“In language teaching arena, there are many notions which are closely interrelated in some way or another, among these, one can refer to instruction and assessment. It is because of this interrelation that any change in one leads to change in another.”(Nasiri, 2020, p. 199). There is a consensus that not only teachers but also students should have a strong role in assessment, technique and procedures. However, the instructor is the sole assessor in traditional approach. A traditional approach is appropriate if students take an objective test but the use of specific assessor in performance assessment such as essays, oral presentation and role playing can contribute to biased evaluations (Matsuno, 2009). Alternative assessments like peer-assessment and self-assessment have been attracting so much attention in an attempt to overcome teacher-assessment limitations (Brown & Hudson, 1998). It has been argued that to train responsible, autonomous, and reflective students, peer-assessment should be used in educational settings (Boud, 1989; Dochy & McDowell, 1997; Falchikov,

1986; Freeman, 1995; Jafarpur, 1991; Somervell, 1993). There are different definitions for peer-assessment. Topping & Ehly (1998, p. 250) defines peer assessment as “an arrangement in which individuals consider the amount, level, value, worth, quality of success of the products or outcomes of learning of peers of similar status.” Peer assessment is a scale tool and a learning cycle that advances skills (Dochy & McDowell, 1997; Sluijsmans, Brand-Gruwel, & van Merriënboer, 2002). A research by Alderson in 1985 (as cited in Pilantana & Thanya, 2011) points out that peer assessment approach is required to improve students understanding and metalinguistic skills related to self-assessment and critical monitoring. McLaughlin & Simpson (2004) proposed that peer-assessment affect learners’ perspectives. Encouraging language learners and involving them in classroom assessment have been demonstrated by many scholars. The main advantages of peer-assessment include providing a chance for learners and helping them develop their English skills. Peer assessment outcomes convert into instructional developments with the purpose of heightening learning, motivation and the students’ confidence (Stiggins, 2008). Motivating learners to be more diligent in their work and enhancing their learning process is the main advantages of peer-assessment procedure (Black, Harrison, & Lee, 2003). According to Reinders & Lázaro (2007), when peer assessment is well organized and applied, it can increase learner’s autonomy and cooperation and enables learners to understand their own learning process and critically evaluate it. Inspiring learners and empowering them to be independent and think better are among the benefits of peer-assessment (Cheng & Warren, 2005; Nilson, 2003; Orsmond, Merry, & Reiling, 1996; Sivan, 2000). A successful classroom atmosphere needs valuable assessment techniques like peer-assessment because such techniques help learners to manage their learning while assessing other students’ performance (Black & Wiliam, 2009). Regarding the disadvantages of peer assessment, (McDowell, 1995) argues that in organizational phase some weaker aspects like anxiety, cheating, and time limitation may emerge. As a result, it is probable to see a mismatch between marks and learning. Even following planning and instruction, students may feel anxious (at least) at the beginning of peer assessment process (Falchikov, 2013; Topping & Ehly, 1998). Poor reliability and validity is the other shortcoming in peer assessment process (Bostock, 2000; Matsuno, 2009; White, 2009). Topping & Ehly (1998) examined the role of peer-assessment in classroom outcomes and showed that it plays a significant role and leads to positive outcomes. Patri (2002) investigated the relationship between peer-assessment and classroom quality in a correlational descriptive study. The findings indicates that, classroom quality would increase if teachers use peer-assessment technique for oral presentation skills. Birjandi and Siyyari’s study (2010) (as cited in Fazel, 2015) studied the effect of self and peer-assessment on learner’s writing production variability and results showed that peer assessment was more successful than self-assessment. Likewise, another study has been conducted on the writing ability of Iranian university students and showed positive outcomes of peer-assessment in this regard (Birjandi & Hadidi Tamjid, 2012). Zarei and Sayar Mahdavi (2014) explored the effect of peer and teacher assessment on Iranian intermediate-level EFL learners’ lexical and grammatical writing accuracy. The findings gave confirmation to the positive and significant effect of peer assessment on lexical and grammatical writing accuracy of the participants.

However, in spite of a good deal of debate concerning the advantages of peer evaluation in the language teaching still, peer evaluation is not commonly used in many EFL context especially Iran. In other words, a look at Iranian educational system shows the traditional assessment is still commonly used. More particularly, during exam sessions, all students are silent, they cannot ask any questions, teachers or examiner does not give them any hint or guidance, and their performance is evaluated based on scores. With a view to this, coupled with the positive effects of peer assessment on language learning and empty place of an empirical study on the effect of peer assessment on improvement of oral production of Iranian intermediate EFL learners in the existing literature, the present research is set out to touch this gap.

1.1. Assessment

According to Segers (2006), applying assessment is not just measuring students' performance, but it also covers some other features such as learners' participation, using knowledge and skills, homogenizing learning environment or building knowledge. In another description, Huba and Freed (2000) stated that assessment is a vehicle of collecting information from various sources or situations, discussing them and developing learners' understanding and finding out what the learner could do with such knowledge consequently. Assessment can be described as a procedure to consider how much learners obtained the predicated criteria or instruction process (Gronlund, 1998). According to Wikström (2008), applying modern methods like CLT and more interactive techniques in classroom programs made the evaluation outstanding. As a result, in an interactive atmosphere, the classroom members are more engaged in each process; they work on activities shoulder by shoulder while teacher monitors and helps to keep the process as interactive as possible. The most effective way of supporting learners' learning and encouraging them to go forward is assessment. When the instructor use assessment strategies, the program may be useful for majority of learners. Assessment is an essential and important segment of learning and teaching progress which may be done during the course not only at the end of the course (Brown, 2004).

1.2. Peer-assessment

In peer-assessment technique learners have the opportunity to participate in evaluating or providing feedback to their peers which is the best way to enhance learning process as well as encouraging learners' engagement (Boud & Falchikov, 2007). Peer-assessment is a means of enhancing learners' participation in classroom activities and involving them in teaching and learning process. Using appropriate instruction and monitoring in instructional settings may involve students in summative assessment (Boud & Falchikov, 2007). Peer- assessment is the mechanism by which peers report formally or informally on their results (Casson, 2009). Scholars have highlighted that the basic aspect of using peer-assessment technique in classroom atmosphere is involving learners in their own learning process and carrying out tasks (Gibbs & Simpson, 2005). First of all, feedback should be adjusted according to learners' need based on their performance especially during their learning process (Gibbs &

Simpson, 2005). Cheng and Warren (1997) believed that to obtain the best result through using peer assessment procedure, the learners need to interact with peers. Therefore, using this procedure in the classroom may have some advantages, for example, it might increase learner's social skills when they give comments and suggestions to their peer's production (Tahir, 2012). When the students use peer-evaluation techniques, they would improve writing as well as communication simultaneously. According to Jensen and Fischer (2005), the subjects who participate in the peer-evaluation process had far better progress in writing skills than those in classes which were based on instructor-centered method. Next, it should take into account the unique characteristics of learners (Gibbs & Simpson, 2005). So, this kind of transferring feedback or interaction between students can be useful for different courses and assignments. It is worth to mention that peer-assessment has a great role in motivating students particularly in mental decision making and enhancing their learning (Casson, 2009).

1.3. Research questions

Q1: Does peer assessment have any effect on oral production of Iranian pre-intermediate EFL learners?

Q2: Does gender have any effect on oral production of Iranian pre-intermediate EFL learners?

Q3: Does the interaction between peer assessment and gender have any effect on oral production of Iranian pre-intermediate EFL learners?

1.4. Research Hypotheses

H01: Peer assessment does not have any effect on oral production of Iranian Pre-intermediate EFL learners.

H02: Gender does not have any effect on oral production of Iranian Pre-intermediate EFL learners.

H03: The Interaction between peer and gender does not have any effect on oral production of Iranian Pre- Intermediate EFL students.

2. Design of the study

The present study sought to examine whether using peer-assessment could improve the oral production of EFL learners. In addition, the study intended to consider the effect of gender on improvement of oral production as well as the interaction among peer-assessment and gender. To do this, the true-experimental- design including randomization, pre-test-and post-test was employed. There are three different variables in this study. Peer-assessment and gender are independent variables, whereas oral production is the dependent variable.

2.1. Participants and settings

The participants of the research consisted of 120 EFL male and female learners, they were studying English at pre-intermediate level in Kalam language institute, in Ardabil. The participant's homogeneity was checked through administering a version of Oxford Placement Test (OPT). After this process 80 learners were chosen randomly based on their tests result; they were within the age range of 17 to 35. Then they were randomly divided into four, control and experimental groups based on their gender.

2.2. Material and Instruments

Oxford Placement Test was used for the purpose of sampling. Peer assessment questionnaire adopted from Patri (2002) was the other material used in this investigation. Also, Cambridge oral test was used as pre and post-test in order to consider the effect of treatment in this study.

2.2.1. Oxford placement Test

The Oxford placement test (OPT) is a valid English proficiency test designed by Oxford University Press to provide teachers with a time-saving and reliable methods of determining the learners' level of English. The pen and paper version was used in the present study. This test consists of 53 items, including 50 multiple-choice, 2 graded true-false items for reading and an optional item for writing. The 50 multiple-choice language use items with each question carrying one mark, give a total score of 70 and students are supposed to choose the correct answer from among the alternatives. Table1 shows the structure of the test. Those learners who scored 31-47 (pre-intermediate) were selected for the main phase of the study.

Table 1. The range of scores for different proficiency levels

	Total	Elementary	Pre-Intermediate	Intermediate
Grammar& Vocabulary	50	0-20	21-30	31+
Reading	10	0-4	5-7	8 +
Writing	10	0-4	5-7	8+

2.2.2. Cambridge oral test

The Cambridge Oral Test (COT) is one of the standard formats of speaking and trustable ways to test overall English proficiency. The formulated questions intended to display a conversation between two people who encounter for the first time. The test includes 36 questions in total. Each item receives a score from 0 to 3. And the questions should be selected in such a way that conversation flows as naturally as possible. Based on the test rules and regulations the researcher asked three questions from question bank 1 and recorded them but did not score this performance. In this way during the question and answering process if a student had a high communicative competence the researcher shifted to questions of bank 3. Also, in the remaining time, the researcher asked follow up questions to discover more concerning students background knowledge. The allocated time was 5 minutes for each

student. Furthermore, the scoring process was done in two phases, first by the researcher then with a colleague to ensure reliability.

Table 2. The oral placement test (Cambridge university)

Criteria	Score
The student doesn't understand the question, even if it is repeated.	0
Evidence of simple understanding, but short (word or phrase only) and/or inaccurate answers, frequent hesitation, limited range of vocabulary with little or no evidence of ability to extend answers, and pronunciation which seriously impedes understanding.	1
Clear evidence of comprehension and ability to form longer answers when appropriate, with non-impeding errors, only occasional hesitation, and ability to self-correct if necessary, but complex vocabulary and grammatical structures are avoided except in obviously well-reshared utterances. Pronunciation is generally intelligible, with limited strain on the listener.	2
Evidence of full understanding, with complete answers and hesitation occurring only naturally whilst an appropriate response is being formulated. Some non-impeding inaccuracy and unnatural language choices, but an obvious ability to extend answers and use complex vocabulary and grammatical structures when appropriate. Fully intelligible pronunciation with no significant strain on the listener.	3

2.2.3. Peer assessment questionnaire

Peer assessment questionnaire is the other instrument used in this investigation was adopted from Patri (2002). The questionnaire consists of fourteen questions divided into four categories: a) organization and content b) use of presentation c) manner d) Interaction with audience. Grading process was done based on a 5-point Linkert scale in which the orders are as follows: 1. Poor 2. Unsatisfactory 3. Satisfactory 4. Good and 5. Excellent.

2.2.4. Textbook

The researcher used For and Against 1 by Gilian Flaherty. The book includes 15 units which covering four language skills. The central objectives of the book are to support students obtain confidence and express their thoughts, opinion and ideas. As a result, they are supposed to gradually develop speaking skills and discussion.

2.3. Procedure

To collect the data, first, the sample was homogenized through the OPT test. Then, eighty learners were chosen based on the result of the OPT. Next, the selected sample was divided into four groups namely, two experimental groups and two control groups. In the next step, the Cambridge oral test was administered in the four groups as the pre-test. Then, the treatment period started which included sixteen one-hour class sessions which were held twice a week. In the experimental groups, the researcher devoted the first three sessions of the treatment period to training the learners, familiarizing them with the key elements of a good oral presentation and using some practical activities to help them stay motivated and not to get bored. The course syllabus used in the mentioned sessions contained making oral presentation with main purpose of organization, content, language use, manner and interaction with the audience. In the remaining thirteen sessions, the groups were divided into two halves: the first half was given subject A and the second half subject B, which were

selected by the researcher. Also, each class was divided into groups of five according to the class size. Participants were asked to change their groups weekly to enhance learners' motivation and help them experience new perspectives. The allocated time for each presentation was three minutes. In each class session, during the presentations, the members of each group took note and filled the questionnaire. Upon the completion of the presentation, the groups compared their evaluations and the teacher explained the differences between the evaluations and commented on them. However, the control groups did not receive any special treatment and were just exposed to traditional mainstream teacher-centred method of teaching oral skill. One week after the end of the class sessions, all the four groups sat the Cambridge oral test as the post-test.

2.4. Data analysis

Descriptive statistics, one-way, and two-way ANOVA were used to analyse the data. One-way ANOVA was run to measure the significant difference between proficiency level prior to the main phase of the study. Finally, a two-way ANOVA was used to examine if there is any improvement as a result of the treatment and gender difference. All the statistics in this study were conducted by applying SPSS 21 software.

3. Results

3.1. Testing Normality of the Data

As the most important assumption for the parametric tests, normality of the collected data was checked. Table.3 displays the skewness and kurtosis statistics and their ratios over the standard errors.

Table 3. Normality test

	N	Min	Max	Mean	STD	Skewness	Kurtosis		
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Control Group (Male)	20	14.00	34.00	22.1500	7.30014	0.196	0.512	-1.780	0.992
Control Group (Female)	20	14.00	33.00	22.0000	6.10436	0.154	0.512	-1.588	0.992
Experimental Group (Female)	20	14.00	33.00	21.4500	5.83524	0.420	0.512	-1.207	0.992
Experimental Group (Female)	20	14.00	36.00	21.5500	6.70016	0.780	0.512	-.407	0.992
Control Group (Male)	20	30.00	42.00	35.0000	3.82512	0.251	0.512	-1.316	0.992
Control Group (Female)	20	28.00	40.00	34.8000	3.50338	-0.092	0.512	-1.040	0.992
Experimental Group (Female)	20	57.00	66.00	59.6500	3.18343	0.714	0.512	-1.215	0.992
Experimental Group (Female)	20	56.00	66.00	59.8000	3.45802	0.561	0.512	-1.412	0.992
Valid N (listwise)	20								
Control Group (Male)									

All the ratios were within ± 1.96 , so the data were considered normally distributed and we are allowed to run two-way analysis of variance (Two-Way ANOVA) as a parametric test.

3.2. Testing the hypothesis of the research

A pre-requisite to any comparison of three independent means is equality of variances. Equality of variances was investigated using Levene’s test. Tab.4 labeled homogeneity of variances shows that the variance in the post-test scores is similar for each of the groups:

Table 4. Levene's test of equality of error variances of groups across gender

F	df1	df2	Sig.
.494	3	76	0.687

*Tests the null hypothesis that the error variance of the dependent variable is equal across groups. a. Design: Intercept + Groups + Gender + Groups * Gender*

Tab.5 Depicts proficiency test of the present paper:

Table 5. Descriptive statistics for the scores on proficiency test

Group	N	Min. Score	Max. Score	Mean Score	SD
Male	60	22	48	43	2.63
Female	60	19	55	42	3.45

Descriptive statistics was run to compare the groups’ means on proficiency test in order to prove that the two groups were homogenous in terms of their general language proficiency prior to main study. Table.5 displays the results of the descriptive statistics for the two groups of male and female participants on the test. The results indicated that the male (M = 43, SD = 2.63) and female (M = 42, SD = 3.45) groups had fairly close means on the proficiency test with signifying a slightly higher mean score among male students.

Table 6: Descriptive statistics for the scores on pre-test across gender

Groups	Dependent Variable: Scores			
	Gender	Mean	Std. Deviation	N
Control	Male	22.8500	7.48862	40
	Female	22.4000	5.99059	40
	Total	22.6250	6.74185	80
Experimental	Male	22.4500	5.98695	40
	Female	22.0250	7.22704	40
	Total	22.2375	6.59736	80
Total	Male	22.6500	6.73945	80
	Female	22.2125	6.59822	80
	Total	22.4313	6.65183	160

Since gender was the basis for discrimination of the nominal variables, there were similar numbers of male and female students in each group. The control group consisted of 40 learners in each group of genders as well as the experimental groups in which males and females were similarly distributed. The descriptive statistics in Tab.6 show that the means scores in control groups are 22 with a slight difference in males mean scores and similarly 22 in the experimental groups with a small difference for male groups. According to the data

obtained in pre-tests, total mean score is 22.62 for control and 22.43 for experimental groups. So, all groups before treatment do not differ from each other on the pre-tests. To find the amount of achievement after treatment, a set of descriptive statistics was performed for the scores on the post-tests in two groups. The results of this test are reported in the following table.

Table 7. Descriptive statistics for the scores on post-test across gender

Groups	Dependent Variable: Scores			
	Gender	Mean	Std. Deviation	N
Control	Male	35.0000	3.82512	20
	Female	34.8000	3.50338	20
	Total	34.9000	3.62187	40
Experimental	Male	59.6500	3.18343	20
	Female	59.8000	3.45802	20
	Total	59.7250	3.28155	40
Total	Male	47.3250	12.95631	40
	Female	47.3000	13.11722	40
	Total	47.3125	12.95425	80

According to Tab.7, the mean and standard deviation (STD) of male learners ($M = 35$, $SD = 3.8$) and the female learners ($M = 34.80$, $SD = 3.5$) in control groups do not differ much from each other on the post-test after treatment. Moreover, Tab.7 reflects that the mean and standard deviation of the male learners in experimental ($M = 59.65$, $SD = 3.1$) and female learners ($M = 59.80$, $SD = 3.45$) on the post-test with fairly close scores. As it is clear from the table, the mean scores of participants in experimental groups are higher than the scores in control groups. It shows that peer assessment made them enjoy higher level of social cognition than monolingual learners.

The actual result of the two-way ANOVA – namely, whether either of the two independent variables (groups and gender) or their interaction is statistically significant – is shown in the *Tests of Between-Subjects Effects* table, as shown below in Tab.8:

Table 8. Two-way ANOVA for the scores on post-test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	12326.237 ^a	3	4108.746	335.426	0.000	0.930
Intercept	179077.812	1	179077.812	1.462E4	0.000	0.995
Groups	12325.613	1	12325.613	1.006E3	0.000	0.930
Gender	0.013	1	0.013	0.001	0.975	0.000
Groups * Gender	0.612	1	0.612	0.050	0.824	0.001
Error	930.950	76	12.249			
Total	192335.000	80				
Corrected Total	13257.187	79				

a. *R Squared = .930 (Adjusted R Squared = .927)*

The particular rows we are interested in are the "Groups", "Gender" and "Groups* Gender" rows. These rows inform us whether our independent variables (groups and gender) and their interaction (groups* gender) have a statistically significant effect on the dependent

variable (oral production). It is important to first look at the " Groups* Gender " interaction as this will determine how we can interpret our results. The "Sig." column shows that there is not a statistically significant interaction at the $p = 0.824$ level. Further, it can be seen from Tab. 8 that there was statistically significant difference between experimental and control groups ($p = .000$), but there was not statistically significant difference between male and female i.e., gender ($p = .975$). It is represented graphically in Fig.1.

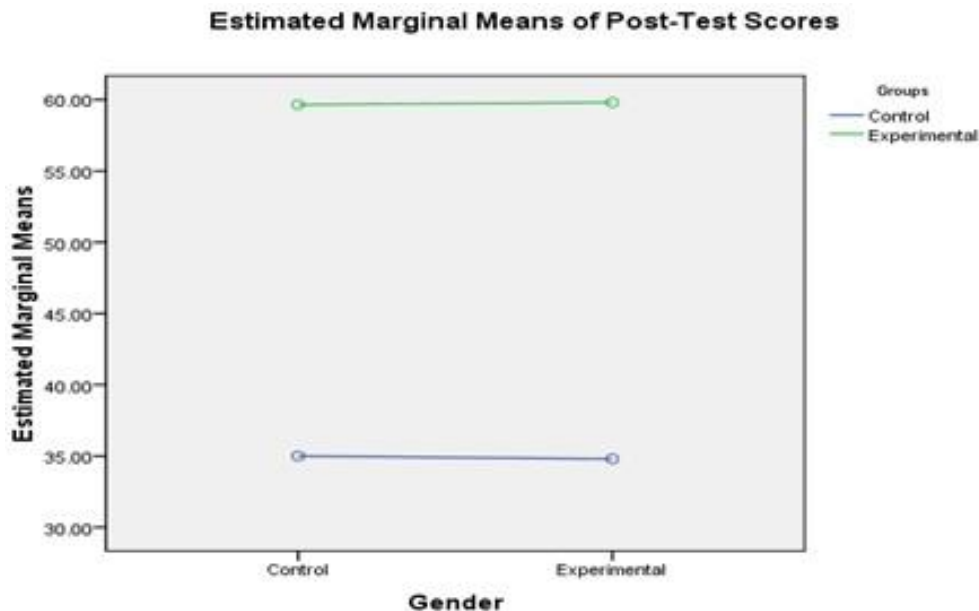


Figure 1. Experimental and Control Group Scores across Gender

4. Discussion

The present study tried to investigate the effect of peer assessment on improvement of Iranian pre-intermediate EFL learner's oral production. To reach the main purpose of the study three questions and hypotheses were structured as following:

1. Does peer-assessment have any effect on oral production of Iranian EFL learners?
2. Does gender have any effect on oral production of Iranian pre-intermediate EFL learners?
3. Does the interaction between peer-assessment and gender have any effect on oral production of Iranian Pre-intermediate EFL learners?

H₀₁: Peer assessment does not have any effect on oral production of Iranian Pre-intermediate EFL learners?

H₀₂: Gender does not have any effect on oral production of Iranian Pre-intermediate EFL learners?

H₀₃: The Interaction between peer and gender does not have any effect on oral production of Iranian Pre- Intermediate EFL students?

The advantages of applying alternative strategies like peer and self-assessment on EFL contexts endorsed by many scholars such as Brown, 2001; Cheng & Warren, 2005; Dyer, 1996; Freeman, 1995; Hughes & Large, 1993; Miller & Ng, 1994; Nilson, 2003.

Topping (2009) believed that using peer assessment provides a number of advantages for the assessors as well as assesses. It is comprised of increased level of time on assignments and practice, and enhance sense of social responsibility. Actively involving learners prompting them to ask intelligent questions, and understanding the assessment process is the main feature of formative peer-assessment mentioned by Topping (2009). Moreover, cognitive and metacognitive awareness achievements may appear in different way like before, after or during the peer-evaluation process. Scruggs and Mastropieri (1998) claimed that employing peer evaluation in various levels such as elementary, high schools, or people with special educational needs or learning disabilities was successful. Different research has proven that peer evaluation has many advantages. In line with this findings, (Alderson, 1985), peer assessment technique might raise learners understanding through the involving them in the process. Also, it may contribute to learner's meta-cognitive improvements which integrate with self-evaluation and critical monitoring. On the other hand, using peer evaluation technique has some limitation such as determining the criteria for peer-evaluation and student's language proficiency. Based on the findings of study conducted by (Cheng & Warren, 2005), students were nervous and uncertain on their skills to assess their peer mate's language proficiency when they asked to evaluate based on the particular language criteria. Furthermore, the results of (Miller & Ng, 1994) revealed that the reliability of peer evaluation went under the question with and student's unsatisfactory comments. They mentioned that the assessors should have especial language skill and prefer to have evaluation by teachers than peers. Generally, it is recommended that peer assessment is not much time consuming but Falchikov (2003) claimed that it is difficult to save time. Particularly in short medium terms as it is requiring well-organization, training and observing the peer-evaluation quality. They mention that the assessors should have especial language skill and prefer to evaluate their performers by teachers than by peers. Kaufman & Schunn (2011) conducted a study to consider the effect of online peer-assessment system SWORD (students writing origin revision work) on the writing assignments. The results of the study showed that pupil believe that applying this technique is unfair and peers are not as sophisticated as instructors to analyse their classmates works. Other research administered by (Nian-shhing, Chen-Wang wei, Kuen-ting We, Lorna Udea, 2008) indicated that applying peer assessment has no significant influence on the reflection levels. In this way, they studied the influence of high-level prompts and peer assessment on online learner's reflection levels. Chiramanee and Chiramanee (2014) conducted a study on peer evaluation of oral English proficiency. According to the findings of the study, learners with varying levels of language ability rate their peers differently and are not similar to teacher's evaluations. This could imply that learner's linguistic proficiency has no bearing on peer evaluation. Yang, Badger, & Yu (2006) conducted a study and considered the effect of using peer and teacher assessment on improvement of writing skill in Chinse university. Based on the findings of the research they found out majority of students were eager to receive peer-assessment suggestions rather than

teacher comments. The present inquiry almost was in line with what Patri (2002) achieved in his study. The researcher considered the influence of peer-feedback on self and peer-assessment of oral skills. And they were selected the participants among 56 first year undergraduate Chinese students whereas the present study chosen the participants among pre-intermediate institute learners and split them into experimental and control based on their gender. Also, Graves (2013) conducted a study and examined the impact of applying peer-review on writing. The findings of the study indicated that at the beginning of the study a large number of participants faced with problem in writing process but as soon as they get into the peer-assessment program gradually made progress and developed their skills. The results of this research confirmed what Birjandi & Hadidi Tamjid (2012) found in their study. They claimed that the students writing skill strongly developed through applying peer assessment technique.

The current study also was in line with a study administered by Williams (1992). In that study, the researcher investigates the effect of using self and peer assessment techniques. The research participants were selected from among the 99-biasness male and female students. Nonetheless, the present study only focused on peer-assessment and examined the effect of utilizing such procedures on oral production of the Iranian EFL learners with regard to their gender.

The outcomes of the present study illustrated that there was statistically significant difference among the experimental and control group and the peer-assessment technique improved the oral production of participants. The results of this study contradicted what Freeman (1995) found in his study. The researcher proposed that there was no significant difference between peer-assessed and staff-assigned grades.

5. Conclusion

The outcomes and analysis of the current study demonstrated that implementing the peer-assessment strategy had positive effect on EFL context and there was a strong difference between the results of pre and post-test of the control and experimental groups. Students were satisfied with the process and developed their skills. Based on the data analysis of the study the first null hypothesis of the inquiry that peer assessment does not have any effect on oral production of Iranian Pre-intermediate EFL learners was rejected. Since the mean scores for the male learners ($M = 59.65$), ($SD = 3.18$) were not significantly different from that the female learners consequently the second hypothesis rejected. Also, further consideration of the results indicated that there was no strong difference among the interaction of the gender and peer assessment.

6. Pedagogical implications

The importance and the necessity of assessment and oral production for language learners cannot be disregarded. Generally, the findings of the study suggested awareness about the positive effects of peer assessment as a significant contribution to academic achievements in

productive skills. Therefore, after acknowledging the importance of peer assessment, teachers and practitioners in EFL should employ such techniques or strategies to teach oral skill. Syllabus designers and materials writers may use the results of this investigation particularly in Iranian atmospheres. The implication of using peer-assessment in this study had significant effect on experimental groups. Accordingly, it is highly suggested to use more assessment practices both in high schools and universities. Fostering students learning and developing motivation is the key elements of peer evaluation in learning and developing autonomous learners. In the theory, peer evaluation can enable learners to actively regulate their own learning. According to (Butler & Winne, 1995) this is a feature of self-regulated learning that allow pupils to improve through input from external sources such as peers. Also, it transforms traditional teacher-centred classes into students-centred one and encouraging learners and teacher to share their responsibility. In addition, if students broaden their awareness of learning habits out of the course and continue after normal instruction, the authorities will be able to make knowledgeable decisions. As the beneficial impact of peer evaluation was observed in this study, teaching aids should construct in such a way to promote peer evaluation techniques and help learners progress.

7. Suggestions for Further Research

The current study may open new directions for further research. The study attempted to explore the effect of peer assessments on EFL pre-intermediate learners. Therefore, the interested scholars may focus on other language skills such as reading and writing. Since the present study used a limited number of participants, future researchers may choose a large number of students. Also, they may select participants from other proficiency levels.

As regard with the type of assessment, another study can consider the different types of assessment in second language classrooms. Researchers can apply other kinds of alternative assessments like portfolio, observation, and interview for the data collection process in teaching different language skills. There is limited research on the integration of group works in language learning classrooms so, it created a need for further study in this area.

Other study can be conducted considering the learners' attitudes toward peer assessment and teacher's assessment. Finally, it can be valuable to interview with some students to gain more insight into their feelings and attitudes towards the process and types of assessment applied in speaking classes.

References

- Alderson, J.(1985). Evaluation: a way of involving the learner. *Lancaster Practical Papers in English Language Education*, Vol. 6, pp. 45–80.
- Birjandi, P., & Hadidi Tamjid, N. (2012). The role of self-, peer and teacher assessment in promoting Iranian EFL learners' writing performance. *Assessment & Evaluation in Higher Education*, Vol. 37(5), pp. 513–533.
- Birjandi, P., & Siyyari, M. (2010). Self-assessment and peer-assessment: A comparative

- study of their effect on writing performance and rating accuracy. *IJAL*, Vol. 13(1), pp. 23–45.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, Vol. 21(1), pp. 5–31.
- Black, P., Harrison, C., & Lee, C. (2003). *Assessment for learning: Putting it into practice*. McGraw-Hill Education (UK).
- Bostock, S. (2000). Student peer assessment. *Learning Technology*, Vol. 5(1), pp. 245–249.
- Boud, D. (1989). The role of self-assessment in student grading. *Assessment in Higher Education*, Vol. 14(1), pp. 20–30.
- Boud, D., & Falchikov, N. (2007). *Rethinking assessment in higher education: Learning for the longer term*. New York: Routledge.
- Brown, D. in press. H. (2004). Language Assessment: Principles and Classroom Practices. *Person Longman*.
- Brown, H. D. (2001). *Teaching by Principles: An Interactive Approach to Language Pedagogy*. San Francisco State University: LongmanLongman.
- Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL quarterly*, Vol. 32(4), pp. 653–675.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of educational research*, Vol. 65(3), pp. 245–281.
- Casson, A. (2009). Assessment in outdoor education. Citeseer.
- Chen, N.-S., Wei, C.-W., Wu, K.-T., & Uden, L. (2009). Effects of high level prompts and peer assessment on online learners' reflection levels. *Computers & Education*, 52(2), 283–291. <https://doi.org/10.1016/J.COMPEDU.2008.08.007>
- Cheng, W., & Warren, M. (1997). Having second thoughts: Student perceptions before and after a peer assessment exercise. *Studies in Higher Education*, Vol. 22(2), pp. 233–239.
- Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language testing*, Vol. 22(1), pp. 93–121.
- Crandell, C. C., & Smaldino, J. J. (2000). Classroom Acoustics for Children With Normal Hearing and With Hearing Impairment. *Language, Speech, and Hearing Services in Schools*, Vol. 31(4), pp. 362–370.
- Dochy, F. J. R. C., & McDowell, L. (1997). Introduction: Assessment as a Tool for Learning. *Studies in educational evaluation*, Vol. 23(4), pp. 279–298.
- Dyer, B. (1996). L1 and L2 composition theories: Hillocks' 'environmental mode 'and task-based language teaching. *ELT journal*, Vol. 50(4), pp. 312–317.
- Falchikov, N. (1986). Product comparisons and process benefits of collaborative peer ground and self assessments. *Assessment & Evaluation in Higher Education*, Vol. 11(2), pp. 146–166.
- Falchikov, N. (2003). *Learning together: Peer tutoring in higher education*. Routledge.

- Falchikov, N. (2013). *Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education*. Routledge.
- Fazel, I. (2015). A step in the right direction: Peer-assessment of oral presentations in an EFL setting. *Reading Matrix: An International Online Journal*, Vol. 15(1), pp. 78–90.
- Freeman, M. (1995). Peer assessment by groups of group work. *Assessment & Evaluation in Higher Education*, Vol. 20(3), pp. 289–300.
- Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and teaching in higher education*, Vol. 1(1), pp. 3–31.
- Graves, R. (2013). Why students struggle with writing: What to do about it. *University Affairs*, Vol. 54(8), pp. 37.
- Gronlund, N. E. (1998). *Assessment of student achievement*. ERIC.
- Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning*. ERIC.
- Hughes, I. E., & Large, B. J. (1993). Staff and peer-group assessment of oral communication skills. *Studies in Higher Education*, Vol. 18(3), pp. 379–385.
- Jafarpur, A. (1991). Can naive EFL learners estimate their own proficiency? *Evaluation & Research in Education*, Vol. 5(3), pp. 145–157.
- Jensen, W., & Fischer, B. (2005). Teaching Technical Writing through Student Peer-Evaluation. *Journal of Technical Writing and Communication*, Vol. 35(1), pp. 95–100.
- Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: their origin and impact on revision work. *Instructional Science*, Vol. 39(3), pp. 387–406.
- Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language testing*, Vol. 26(1), pp. 75–100.
- McDowell, L. (1995). The impact of innovative assessment on student learning. *Innovations in Education and Training International*, Vol. 32(4), pp. 302–313.
- McLaughlin, P., & Simpson, N. (2004). Peer assessment in first year university: How the students feel. *Studies in Educational evaluation*, Vol. 30(2), pp. 135–149.
- Miller, L., & Ng, R. (1994). Peer assessment of oral language proficiency. *Perspectives: working papers of the department of English, City Polytechnic of Hong Kong*, Vol. 6(2), pp. 41–56.
- Nasiri, S. (2020). A Review on Dynamic Assessment (DA) in Iran. *Critical Literary Studies*, Vol. 2(1, Autumn and Winter 2019-2020), pp. 199–211.
- Nilson, L. B. (2003). Improving student peer feedback. *College teaching*, Vol. 51(1), pp. 34–38.
- Orsmond, P., Merry, S., & Reiling, K. (1996). The importance of marking criteria in the use of peer assessment. *Assessment & Evaluation in Higher Education*, Vol. 21(3), pp. 239–250.
- Patri, M. (2002). The influence of peer feedback on self-and peer-assessment of oral skills.

- Language testing*, Vol. 19(2), pp. 109–131.
- Reinders, H., & Lázaro, N. (2007). Innovation in Language Support: The provision of technology in self-access. *Computer Assisted Language Learning*, Vol. 20(2), pp. 117–130.
- Scruggs, T. E., & Mastropieri, M. A. (1998). Tutoring and students with special needs. *Peer-assisted learning*, pp. 165–182.
- Segers, M., Dochy, F., & Cascallar, E. (2006). *Optimising new modes of assessment: In search of qualities and standards* (Vol. 1). Springer.
- Sivan, A. (2000). The implementation of peer assessment: an action research approach. *Assessment in education: Principles, policy & practice*, Vol. 7(2), pp. 193–213.
- Sluijsmans, D. M. A., Brand-Gruwel, S., & van Merriënboer, J. J. G. (2002). Peer assessment training in teacher education: Effects on performance and perceptions. *Assessment & Evaluation in Higher Education*, Vol. 27(5), pp. 443–454.
- Somervell, H. (1993). Issues in assessment, enterprise and higher education: The case for self-peer and collaborative assessment. *Assessment and evaluation in Higher Education*, Vol. 18(3), pp. 221–233.
- Stiggins, R. (2008). A call for the development of balanced assessment systems. *Assessment Manifesto*. Portland, OR: ETS Assessment Training Institute.
- Tahir, I. H. (2012). A Study on Peer Evaluation and Its Influence on College ESL Students. *Procedia - Social and Behavioral Sciences*, Vol. 68, pp. 192–201.
- Topping, K. J. (2009). Peer assessment. *Theory into practice*, Vol. 48(1), pp. 20–27.
- Topping, K., & Ehly, S. (1998). *Peer-assisted learning*. Routledge.
- White, E. (2009). Student perspectives of peer assessment for learning in a public speaking course. *Asian EFL Journal*, Vol. 33(1), pp. 1–36.
- Wikström, N. (2008). Alternative assessment in primary years of international baccalaureate education.
- Williams, E. (1992). Student attitudes towards approaches to learning and assessment. *Assessment and evaluation in higher education*, Vol. 17(1), pp. 45–58.
- Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of second language writing*, Vol. 15(3), pp. 179–200.
- Yinjaroen, P., Chiramanee, T., & Chiramanee, T. (2014). Peer assessment of oral English proficiency.
- Zarei, A. A., & Sayar Mahdavi, A. (2014). The effect of peer and teacher assessment on EFL learners' grammatical and lexical writing accuracy. *Journal of Social Issues and Humanities*, Vol. 2(9), pp. 92–97.