



DIFFERING EFFECTS OF CONFERENCE AND PEER ASSESSMENT ON WRITING SKILL OF EFL LEARNERS

Maryam Manafi Anvar

Islamic Azad University, Bandar Abbas Branch
Department of English, Bandar Abbas Branch, Islamic Azad University, Bandar Abbas, Iran

Azadeh Nemati

Department of English language teaching, Jahrom branch, Islamic Azad University,
Jahrom, Iran

azadehnematiar@yahoo.com

ABSTRACT

This study intended to demonstrate the significant differences between conference and peer evaluation on writing skill of EFL learners in Pooyesh Language House of Bandar Lengeh. To find out the differences, two types of tests (pre- test before treatment and receiving feedback, post- test after treatment and receiving feedback) were given to two groups of students (30 students for conference and 30 students for peer evaluation). The participants were all teen and adult students who were studying at the same level in this Language House. The writing tests were taken from their previous books which they had practiced before. The pre-test was given to students. In conference group, a mark was given to them by their teacher and during the term the students learned how to write through their teacher's comments, but in peer evaluation group, through their classmates' comments, so in the second group the classmates just checked each other's papers and found errors without giving any mark, because they were not familiar with assessment system; it was the teacher' task. In the first group, the teacher was familiar with assessment system, but in peer evaluation group the students found how to check other students' writing as treatment, according to IELTS writing assessment criteria, through their teacher guidance. The post- test was given at the end of the term and previous method was repeated. The results of the study indicated that there was no statistically significant difference between conference and peer evaluation, so the research questions were rejected. The results gained from the two variables were analyzed according to quasi experimental design.

KEYWORDS: Assessment; Conference Assessment; Peer Assessment;

INTRODUCTION

There are different alternatives in assessment and testing but the current study is going to compare the differences between conference and peer evaluation. Commenting on papers is an informal writing evaluation. Through comments, teachers can develop students' writing abilities. According to Matsuno (2009), the primary goal of commenting on papers is for "students to see what works best and least well in the draft so that revision can take place". Comments can come in the form of marginal or terminal comments. Marginal comments are written in the margin and can be comments about content or mechanics. Terminal comments are final comments issued at the end of a paper that direct students in how to improve their drafts or their next papers and revise them (Messick, 1989). "Feedback plays an important role in development of writing, but only when, students and teachers create a partnership through feedback— teachers should engage their students by treating them as a useful scholars, offering instruction" Messick (1989).

Peer-evaluation is based on self-evaluation and provides an opportunity for students to compare their self-evaluation results with those of others. Often, the results of self and peer-evaluation are considered at the same time (Papinczak et al., 2007). Populated use of peer-evaluation and confirming its influence on group projects has been reported in different studies. In peer evaluation, students' belief in validity of the evaluation and privacy of their personal claims are definitely important, especially in writing. The score of each student may change according to the evaluation of her/his peers (Kench, Field, Agudera, & Gill, 2009). Studies demonstrated that peer-evaluation scores are significantly higher than teacher-evaluation score; on the other hand, scores based on self-evaluation is less accurate than peer-evaluation (Papinczak et al., 2007). Meanwhile, existing research does not support the validity of teacher evaluation to measure students' knowledge. (Cunnington, 2001; Neville, 1999; Whitfield & Xie, 2002). Other competencies of evaluation could be mixed with peer and self-evaluation results to increase its accuracy (Sluijsmans, Dochy, & Moerkerke, 1998).

According to Liu and Hansen (2002), from both cognitive and psycholinguistic perspectives, the four theoretical stances which support the use of peer response activities in the writing classroom are "process writing theory, collaborative learning theory, Vygotsky's Zone of Proximal Development (ZPD), and interaction and second language acquisition" (p. 2). Research based on these theoretical stances has provided substantial evidence that peer response activities, in fact, help second language learners develop their L2 writing abilities. Willmot and Crawford (2004) stated, "Peer review is used to moderate the tutor awarded marks" (p. 17-21). It's a good idea to apply peer evaluation in team works and also group projects. Teachers can give each projects a final mark but also need to determine what grade each individual in the group deserves (Brown, 1998). By the way, Brown and Pendelbury (1992) stated, "Our use of peer review in the Personal Effectiveness module attempts to encourage participation in teamwork, improve group dynamics" (p. 17).

The present study aimed to explore the relationship between conference and peer evaluation among the EFL learners in institutes of Bandar Lengeh.



LITERATURE REVIEW

Generally, investigating the effectiveness of different methods of assessing writing in EFL classes has been one of the important and essential aims between researchers in Iran (Elahinia, 2004; Nezakatgoo, 2005; Javaherbakhsh, 2010), but estimating the popularity of those methods has never gained this much attention. By conferencing we can assess the student's oral production or writing performance and try to find the student's learning styles (Pond, 1995). While through peer assessment, students listen to or just check and review one another student's writing (Dochy et al., 1999). Stecker (2005) reported that peer evaluation should be used for every writing in language classrooms. Shepard (2001) asserted that when students are working with their peers, it is enjoyable for them, they also learn how to handle language better through well-structured group project and assessment. Examining and trying different methods of assessing and evaluating writing is limited to alternative assessment methods in Iran and they have been the focus of attention among Iranian EFL researchers in recent years. Many studies have focused and emphasized on the effect of using portfolios, peer/self-assessment, and journals on improving students' writing skill (Elahinia, 2004; Nezakatgoo, 2005; Javaherbakhsh, 2010; Ghorchaei, Tavakoli, & Nejad Ansari, 2010; Moradan & Hedayati, 2011; Sharifi & Hassaskhah, 2011; Mosmery & Barzegar, 2015). Elahinia (2004) demonstrated significant correlation of portfolio and conference assessment in improving students' writing proficiency in a quasi-experimental design. She also reported that students had a positive attitude about portfolio assessment and it's something enjoyable for them. In another quasi-experimental design, Nezakatgoo (2005) found that portfolio-based assessment had a significant and positive effect on improving students' writing skill, too. In Javaherbakhsh (2010), using self, peer and conference assessment in improving writing had significant effect in comparison with a group with traditional methods. Ghorchaei et al., (2010) also claimed that portfolio and peer assessment had a significant effect on improving students' writing when we want to focus on vocabulary, elaboration, and organization. Moreover, Moradan and Hedayati (2011) reported that intermediate students who received conferencing and portfolio-based assessment performed significantly better than those who received scores without any further feedback. Sharifi and Hassaskhah (2011) and Mosmery and Barzegar (2015) were among other researchers who showed that portfolios and self- and peer assessment are effective ways to help students write better assignments.

Although writing evaluation is the most difficult task of the teachers, it's the most important and valuable part of the teachers' job to help each individual in classes (Matsuno, 2009). While assessment refers to a collection of data, or information that can help not only the teachers but also the learners. On the other hand evaluation means the product of assessment (Messick, 1989). The teachers must consider assessment as a guide to evaluate the product of efforts and the progress of their students. Then they can give grade, letter or score to students very easily. Falchikov (2001) defined peer [assessment] evaluation as the process that each group of individuals rate and grade their peers. Peer evaluation is a broad topic and can be qualitative in nature; formative or summative; and administered with self-assessment (Somervell, 1993). Peer assessment positively affect students' motivation (Das et al., 1998). However, literature suggested that a vast majority of self and peer assessment research has been carried out in disciplines. The most important difficulty of peer evaluation is the validity of evaluation at an acceptable level (Dochy et al., 1999). The most important way of ensuring this is to prepare clear and understandable criteria and standards, to make evaluation by hiding the names or to use more than one peer evaluation for one study (Falchikov, 2001). The success of the evaluation is closely related with how much the students accept criteria and process. It will be useful, if lecturers make student active in this process.

RESEARCH QUESTIONS

To this end, the following research questions were addressed:

1. Is there any statistically significant difference between conference and peer evaluation regarding their effects on EFL learners writing proficiency?
2. Is there any statistically significant difference between conference and peer evaluation regarding their effects on EFL learners writing proficiency among male and female?

METHODOLOGY

Participants

In this study, there were two variables and also two groups of participants, some participants in conference and some others in peer evaluation. There were 12 male and 18 female students in conference group. All of them were intermediate students in Pooyesh Language House of Bandar Lengeh. Also, there were 13 male and 17 female students in peer evaluation group. They were high school students and adults who were studying in Pooyesh Language House of Bandar Lengeh, too. All of them were intermediate students.

Instruments

To determine whether conference evaluation had any correlation with peer assessment, three instruments were used in this study. First, a pre-test and a post-test of writing. In order to investigate the students' level of English writing proficiency these two tests were designed. The topics of writing tests were taken from their own book which they had studied before and all of the students were familiar with them. Also another type of instrument was prepared to have more reliable assessment in



students' writing. The reason for selecting the last instrument in the present study is that IELTS is an international and well known test all around the world.

IELTS writing assessment criteria

Providing pre-determined standards and criteria of students' performance are very essential during assessment process. Without a clear understanding of the subject aims, students have a bigger chance of being confused and wasting time trying to find what it is we want them to learn and understand. To tell the truth, writing assessment is really difficult and confusing and maybe different from one teacher to another. Because of that, IELTS writing assessment criteria was used. According to these criteria, the researcher could assess the writing tests easily, accurately and academically. When you plan your teaching method, you can make decision on what you think is the best way to learn a specific knowledge or skill. If you show these goals to your students, then transfer them to their writing skill, it would be really helpful in assisting your students' learning. By the way, these criteria can be used in different kinds of students' performance such as speaking, etc. But, the main question is that how many criteria are needed to show us the problematic areas in students' writing. It's necessary that the criteria should link with scores and grades given to students and learners. The researcher must investigate how IELTS writing papers are marked, so you can help your students understand what the examiners are exactly looking for. Also, IELTS assessment criteria can be used to help students of different levels prepare for the test in the same class.

Procedure

In conference group, the students were asked to write a paragraph on pre-test about the topic given to them. After collecting the papers, the teacher wrote his/ her comments on students' paper which were then given to them, and according to mentioned criteria, a mark out of 20 was given to them (1st session). After receiving this written feedback, a short time around 10 minutes was given to the students to read the comments and ask any questions, identify their problems or understand clarity about what their teacher had written (2nd session). The students were then asked to rewrite their paragraphs according to the received written comments from their teacher and bring them back to the class (3rd session). Then, the papers were collected by the teacher and put in a file for next analysis.

This method was repeated during the term about different topics as treatment. During the term, the treatment considered conference in this project, checked students' writing exercises by the teacher's comments, through 15 sessions of their English class to improve their writing skill. It means that the teacher taught them how to write. The suitable place to do this was classroom. All of the writing exercises and test were taken from their own books which were taught in the institute. It is important to know that pre- tests were the same in both of selected groups. Also, another kind of test was given to the students as a post- tests at the end of the term that it was were the same in both of the selected groups, too. It was given to students, and after checking the papers by teacher, a mark out of 20 was given to them, and previous the way was used again so the result of two tests were compared to each other to see the results in this group.

Also, the students in peer evaluation were asked to write their paragraphs on the pre- test. Next, the teacher collected the papers, deleted the students' names, and chose a coded number to each paper in order to stop any prejudgments in assessing the papers and giving comments and ideas. Then, he/she distributed them among the students and asked them to write their comments according to the mentioned criteria (the 1st session). Then each student read the other student's paper and provided his/her response to that, using the different questions prepared by the researchers on a draft.

Next session, all the students brought the papers and the written comments on each back to the class and submitted them to the teacher. The mark out of 20 was given to them by their teacher. The teacher attached the deleted names to the papers according to the coded numbers, gave them back to the writers, and asked them to revise their first drafts (the 2nd session). Finally, each student used this feedback to rewrite his/her paper and gave it back to the teacher (the 3rd session).

The papers were collected by the teacher in a separate file. Each student should just find the errors and mistakes of one another student, but the mark should be given by the teacher to them since they don't know how to give score; it was the teachers' task. Because of their being unfamiliar with assessment system, the teacher asked them to check their classmates' writing exercises during the term as treatment according to the mentioned criteria through 15 sessions, and the previous method was used during the term.

This method was repeated during the term. The suitable place to do this was classroom. Also, another kind of test was given to the students as a post- tests and checked by classmates. All the tests were given the mark out of 20 by the teacher in both of groups. The results of two tests were compared to each other to see the results in this group.

Data analysis

After collecting the data, the pre- tests and post- tests of both of the selected groups were scored based on the IELTS Writing Assessment Criteria as mentioned before. There were 10 items and criteria and the total score was 20. The statistical analyses were conducted by using the Statistical Package for Social Sciences (SPSS) in order to investigate whether there was a significant difference between conference and peer evaluation. After assessing and comparing the scores of two tests related to two groups of students and using the quasi experimental, design the data were analyzed in next chapter. The main purpose is to know the real difference between conference and peer evaluation and to see which one is more effective. In addition, after giving the post tests to students the aim was that which group had bigger change.



RESULTS AND DISCUSSIONS

Descriptive statistics

Table 1: Statistics of the IELTS writing scores

Gender	Group	Test	N	Mean	Std. Dev.	Skewness	Kurtosis	Min	Max
Male	Peer-evaluation	Pre-score	13	14.69	2.594	-0.581	-1.026	10	18
		Post-score	13	15.92	2.813	-0.492	-1.209	11	19
	Conference	Pre-score	12	13.92	1.505	-0.408	-0.111	11	16
		Post-score	12	14.58	2.151	-0.268	-0.806	11	18
	Total	Pre-score	25	14.32	2.135	-0.260	-0.738	10	18
		Post-score	25	15.28	2.558	-0.161	-1.110	11	19
Female	Peer-evaluation	Pre-score	17	17.18	1.551	0.124	-0.961	15	20
		Post-score	17	18.12	1.219	-0.251	-0.642	16	20
	Conference	Pre-score	18	17.72	1.447	-0.108	-0.806	15	20
		Post-score	18	17.89	1.906	-0.510	-0.809	14	20
	Total	Pre-score	35	17.46	1.502	-0.023	-0.946	15	20
		Post-score	35	18.00	1.590	-0.558	-0.311	14	20
Total	Peer-evaluation	Pre-score	30	16.10	2.383	-0.897	0.653	10	20
		Post-score	30	17.17	2.306	-1.249	1.000	11	20
	Conference	Pre-score	30	16.20	2.384	-0.287	-0.633	11	20
		Post-score	30	16.57	2.569	-0.409	-0.529	11	20
	Total	Pre-score	60	16.15	2.364	-0.576	-0.068	10	20



		Post-score	60	16.87	2.439	-0.768	-0.142	11	20
--	--	------------	----	-------	-------	--------	--------	----	----

The score mean of the peer-evaluation group was 16.1 in the pre-test and 17.17 in the post-test, which shows an increase in scores. On the other hand, the writing score mean of the conference group increased moderately from 16.20 in pre-test to 16.57 in the post-test.

Inferential statistics

Table 2: Kolmogorov-Smirnov Test of normality for WTC and EI components

Gender	Group	Test	Z	Sig.
Male	Peer-evaluation	Pre-score	0.834	0.490
		Post-score	0.676	0.750
	Conference	Pre-score	0.654	0.786
		Post-score	0.843	0.476
Female	Peer-evaluation	Pre-score	0.551	0.922
		Post-score	0.731	0.660
	Conference	Pre-score	0.614	0.845
		Post-score	0.698	0.715

The above Table shows results for Kolmogorov-Smirnov Test. Since p-values were greater than 0.05 ($p > 0.05$) for all variables (scores in pre- and post-test) and in all groups (peer-evaluation, conference, male and female), the statistics were not significant which means that the distributions were normal.

Investigating of the research questions

To answer the first question and to find the effect of treatment on improving scores, instead of a simple comparison between two groups in the post-test, the analysis of covariance (ANCOVA) was used as reported in Table 3.

Table 3: The ANCOVA for the effect of treatment on writing scores

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Pre-test	206.899	1	206.899	85.067	0.000
Group	6.917	1	6.917	2.844	0.097
Error	138.634	57	2.432		
Total	17420.000	60			

Firstly it showed that there was a significant ($p < 0.001$) linear relationship between the pre-test and the post-test scores. Thus it was legitimate to consider the pre-test as a moderator variable in the analysis to incorporate individual differences existed before training. The second result was the main effect of the group which was not significant on scores in the post-test after controlling for the effect of the pre-test ($F = 2.844$, $df = 1, 57$, $P = 0.097 > 0.05$). This implies that scores did not differ



significantly between two groups. This was the main result of ANCOVA which shows that this research question was rejected.

The second research question concerned the gender of the students. As the previous question, the ANCOVA was used to answer this question. In this case, the comparison of peer-evaluation and conference groups was done by two ANCOVA tests among males and females separately.

Males: Table 4: The ANCOVA for the effect of treatment on writing scores among males

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Pre-test	77.112	1	77.112	24.684	0.000
Group	2.764	1	2.764	0.885	0.357
Error	68.727	22	3.124		
Total	5994.000	25			

Firstly it showed that there was a significant ($p < 0.001$) linear relationship between the pre-test and the post-test scores. Thus it was legitimate to consider the pre-test as a covariate variable in the analysis to incorporate individual differences existed before training. The second result was the main effect of the group which was not significant on scores in the post-test after controlling for the effect of the pre-test ($F = 0.885$, $df = 1, 22$, $P = 0.357 > 0.05$). This implies that scores did not differ significantly between two groups among males. This was the main result of ANCOVA which demonstrates that this research question about males did not accept.

-Females: Table 5: The ANCOVA for the effect of treatment on writing scores among females

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Pre-test	21.697	1	21.697	10.875	0.002
Group	2.320	1	2.320	1.163	0.289
Error	63.846	32	1.995		
Total	11426.000	35			

Firstly it showed that there was a significant ($p < 0.001$) linear relationship between the pre-test and the post-test scores. So the pre-test can be considered as a covariate variable in the analysis to incorporate individual differences existed before training. The second result was the main effect of the group which was not significant on scores in the post-test after controlling for the effect of the pre-test ($F = 1.163$, $df = 1, 32$, $P = 0.289 > 0.05$). This implies that scores did not differ significantly between two groups among females. This was the main result of ANCOVA which shows that the research question for females did not accept.

Discussion

Two research questions were focused on and answered based on the findings of the study. About the first research question we have found that no significant correlation exists between teacher assessments with peer evaluation, so there are mixed reviews about this result.



Burnett and Cavaye (1980) showed that there is a weak relationship between teacher assessment and peer-evaluation. It's interesting to know that this idea is relevant to the results of the current study. They also asserted that may be the differences between the results of this study and others is related to the method of students and teacher evaluation or it refers to the statistical method which was used.

Although, Papinczak et al. (2007) claimed that peer-evaluation scores are usually higher than teacher evaluation scores. Also, Machado et al. (2008) found the same results. While, most of the time, there was a significant difference between teacher and other types of evaluation, because the students don't know how to evaluate as well as their teachers especially in self and peer evaluation. Machado et al. (2008) demonstrated that there was always a positive relationship between self and peer evaluation but this correlation was not found between peer- and teacher evaluation grades which is related to the results of this study. Also, about the second research question we have found no significant correlation between teacher assessment with peer evaluation among males and females. So, there are mixed reviews about this result, too.

The question of identifying and understanding possible differences in linguistic styles and forms between males and females has confused linguistic researchers for decades (Trudgill, 1972; Lakoff, 1975; Labov, 1990; Coates, 1998).

Different researches had showed phonological and pragmatic differences between male and female language use in speech (Trudgill 1972; Key 1975; Holmes 1990; Labov 1990; Eckert 1997), informal writings such as student compositions and essays in their classroom (Mulac et al., 1990; Mulac & Lundell, 1994) and also electronic messaging such as sending e-mails, chatting or texting (Herring, 1996).

Some authors such as (Simkins-Bullock & Wildman, 1991) believed that there is no differences between male and female writing styles and performances especially in formal contexts, although some differences was seen in informal texts. Burman and colleagues (2008) demonstrated that some areas of the brain which worked with language, worked harder in girls than boys during language tasks, so it showed that women were really sensitive while they were learning, writing, speaking, etc. It is obvious that girls and boys use different parts of their brain when they are performing different tasks and abilities.

CONCLUSION

Based on the findings of the present study, it can be concluded that there is no statistically significance relationship between teacher assessment and peer evaluation among EFL learners in Pooyesh Language House of Bandar Lengeh, while the existence of relationship and differences between these two components was mentioned in many other researches. So, it seems that peer evaluation is not good substitute for teacher evaluation since the students are not familiar with the assessment and grading system.

As mentioned before, Burnett and Cavaye (1980) showed that there was a weak relationship between teacher assessment and peer-evaluation, but Eva (2001) showed some advantages and benefits of peer-evaluation including long interaction between peers to receive feedbacks, and preparing opportunities to evaluate communication skills, self-learning and self- evaluation which are normally different from traditional evaluation forms. Also, it was indicated that when the learners were directly involved in assessing their work, it would be highly effective in learning process.

In addition, there are other benefits of peer-evaluation like meta-cognitive skills improvement (Ballantyne et al., 2002), and increasing the understanding of the selected subjects (Papinczak et al., 2007). They also asserted that peer evaluation was a kind of powerful tool to assess, and it helped participants to receive useful, constructive and accurate feedbacks.

Limitation of the study

Like others studies, this study included several limitations. It was more reliable to work on all students of country who were studying English as a foreign language. But, they were not available, and it was really time consuming. One of the limitations could be preparing a suitable condition in which honesty could be observed. Accomplishing a good conference and peer evaluation is really time consuming. The number of the students in this institution were limited and also each term there were just 2 or 3 intermediate levels in that. Also, in peer evaluation students are not so fair to evaluate their classmates and they are not familiar with evaluation system, too.

REFERENCES

1. Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education*, 27(5), 427-441.
2. Brown, S. (1998). *Peer assessment in practice*. Birmingham: SEDA.
3. Brown, G., & Pendlebury, M. (1992). *Assessing active learning*. Sheffield:
4. CVCP/USDU.
5. Burman, A.B., Booth, J.R.R. & Bitan, T. (2008). Sex differences in neural processing of language among children. *Neuropsychological*, 46(5), 1349-1382.



6. Burnett, W., & Cavaye, G. (1980). Peer assessment by fifth year students of surgery. *Assessment in Higher Education*, 5(3), 273-278.
7. Coates, J. (ed.) (1998). *Language and Gender: A Reader*. Oxford: Blackwell.
8. Cunnington, J. (2001). Evolution of student evaluation in the McMaster MD program. *Pedagogue*, 10(2), 1-9.
9. Das, M., Mpofu, Dunn, & Lanphear. (1998). Self and tutor evaluations in problem-based learning tutorials: is there a relationship? *Medical Education*, 32(4), 411-418.
10. Dochy, F., Segers, M. & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: a review. *Studies in Higher Education*, 24(2), 331-350.
11. Eckert, P. (1997). Gender and sociolinguistic variation. *Readings in Language and Gender*, 2(1), 64-75.
12. Elahinia, H. (2004). *Assessment of writing through portfolios and achievement tests*.
13. Tehran, Iran: Unpublished Masteral thesis, Teacher Training University.
14. Eva, K. W. (2001). Assessing tutorial-based assessment. *Adv Health Sci Educ Theory Pract*, 6(3), 243-257.
15. Falchikov, N. (2001). *Learning together: peer tutoring in higher education*. London: Routledge- Falmer.
16. Ghoorchaiei, B., Tavakoli, M., & Nejad Ansari, D. (2010). The impact of portfolios assessment on Iranian EFL students' essay writing: A process- oriented approach. *Journal of Language Studies*, 10(2), 35-51.
17. Herring, S. (1996). Two variants of an electronic message schema. *Computer-dated Communication: Linguistic, Social and Cross-Cultural Perspectives*, 2(1), 81-106.
18. Holmes, J. (1990). Hedges and boosters in women's and men's speech. *Language & Communication*, 10(3), 22-25.
19. Javaherbakhsh, M. R. (2010). The impact of self-assessment on Iranian EFL learners' writing skill. *English Language Teaching*, 3(2), 213-218.
20. Kench, P. L., Field, N., Agudera, M., & Gill, M. (2009). Peer assessment of individual contributions to a group project: Student perceptions. *Radiography*, 15(2), 158-165.
21. Key, M. R. (1975). *Male/female Language*. Metuchen: Scarecrow Press.
22. Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change*.
23. Lakoff, R. T. (1975). *Language and women's place*. New York: Harper Colophon Books.
24. Liu, J., & Hansen, J. G. (2002). *Peer response in second language writing classrooms*. Ann Arbor: University of Michigan Press.
25. Machado, J. L., Machado, V. M., Grec, W., Bollela, V. R., & Vieira, J. E. (2008). Self- and peer assessment may not be an accurate measure of PBL tutorial process. *BMC Med Educ*, 5(8).
26. Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, 6(1), 75-80.
27. Messick, S. (1989). Validity. *Educational measurement*, 2 (3), 13- 103.
28. Moradan, A., & Hedayati, N. (2011). The impact of portfolios and conferencing on Iranian EFL writing skill. *Journal of English Language Teaching and Learning*, 8, 115-141.
29. 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.
30. Mulac, A. & T. L. Lundell, (1994). Effects of gender-linked language differences in adults' written discourse: Multivariate tests of language effects. *Language & Communication*, 14(3).
31. Mulac, A., L. B. Studley & S. Blau, (1990). The gender-linked language effect in primary and secondary students' impromptu essays, *Sex Roles*, 23, 9-10.
32. Neville, A. J. (1999). The tutor in small-group problem-based learning: Teacher?



38. Facilitator? Evaluator? *Med Teach*, 21, 393-401.
39. Nezakatgoo, B. (2005). *The effects of writing and portfolio on final examination scores and mastering mechanics of writing of EFL students*. Tehran, Iran: Unpublished Masteral thesis, Allame Tabtba'i University.
40. Papinczak, T., Young, L., Groves, M., & Haynes, M. (2007). An analysis of peer, self, and tutor assessment in problem-based learning tutorials. *Med Teach*, 29(5), 122-132.
41. Pond, K., Ui-Haq, R. & Wade, W. (1995) Peer review: a precursor to peer assessment. *Innovations in Education and Training International*, 32(2), 314-323.
42. Sharifi, A., & Hassaskhah, J. (2011). *The role of portfolios assessment and reflection on process writing*.
43. Shepard, L. A. (2001). The role of classroom assessment in teaching and learning. *Handbook of research on teaching*, 4(3), 1066–1101.
44. Simkins-Bullock, J. A. & B. G. Wildman, (1991). An investigation into the relationship between gender and language. *Sex Roles*, 24-32.
45. Sluijsmans, D., Dochy, F., & Moerkerke, G. (1998). Creating a Learning environment by Using Self-, Peer- and Co-Assessment. *Learning Environments Research*, 1(3), 293-319.
46. Somervell, H. (1993). Issues in assessment, enterprise and higher education: The case
47. self, peer and collaborative assessment. *Assessment and Evaluation in Higher*
48. *Education*, 18(3), 221-233.
49. Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools*, 42(2), 795–819.
50. Trudgill, P. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society*, 1(2).
51. Whitfield, C. F., & Xie, S. X. (2002). Correlation of problem-based learning facilitators' scores with student performance on written exams. *Adv Health Sci Educ Theory Pract*, 7(1), 41-51.
52. Willmot, P., & Crawford, A. (2004). *Proceedings of the international conference on engineering education*. Los Angeles: Longman.