The Effect of Peer, Self, and Traditional Assessment on Iranian EFL Learners' L2 Reading Comprehension

Nasrin Shams

PhD Candidate, University of Isfahan, Iran

Mansoor Tavakoli

Associate Professor, University of Isfahan, Iran

Abstract

This study investigated the impact of peer, self, and traditional (or teacher) assessment on improving EFL learners' reading comprehension. To this aim, 77 Iranian students from a private institute were selected as homogeneous from a population pool of 102 volunteers based on their performance on a standard English proficiency test (Nelson, 2001). They were randomly divided into three experimental groups and subsequently exposed to the research treatment. The three groups received peer, self, and traditional assessment on second language (L2) reading comprehension. Then, the reading comprehension achievement test was given to the students in the three groups to find out their reading comprehension ability after the treatment. Statistical analyses of the results revealed that the peer-assessment group significantly outperformed the traditional assessment group in terms of L2 reading comprehension. The results also showed that that there exist no meaningful differences in the performance of the other two groups on comprehension measures. Hence, the findings of this study indicated that utilizing peer-assessment can be influential in language learning in general and L2 reading comprehension in particular. Results may also have important implications for foreign language syllabus designers and language instructors as well.

Key words: peer-assessment, self-assessment, traditional assessment, L2 reading comprehension

INTRODUCTION

An essential feature of education is assessment and the significance and popularity of student-oriented learning demand alternative techniques of assessment to evaluate teaching and learning. Assessment sets the agenda more persuasively than any syllabus or course outline and it is "one of the most significant influences the students' experience of higher education and all that they gain from it" (for more details see Boud & Associates, 2010, p. 1). In recent years, assessment has generally been seen as one of the key challenges in the field of learning. Assessment, in the broad sense, means "any methods used to better understand the current knowledge

---

Correspondence: Nasrin Shams, Department of English, Faculty of Foreign Languages, University of Isfahan, Isfahan, Iran. Email: shams.nasrin@yahoo.com

that a student possesses" (Collins & O'Brien, 2003, p. 29). According to Crooks (2001), assessment is "any process that provides information about the thinking, achievement or progress of students" (p. 1).

Because assessment is important in teaching and learning, every teacher should assess his/her students' learning regularly. Some of the methods which teachers use to measure their students' learning are paper and pencil tests, oral presentations, standardized tests, and question-and-answer activities. Therefore, teachers spend a great deal of their class time engaged in one type of assessment or another (Stiggins, 2001). On the other hand, assessment of students entails using a well organized system, namely tests, to make judgments about the students' achievement (Gronlund & Linn, 1990). While this type of assessment is a mainstay of educational programs educators and critics from various backgrounds have raised a number of concerns about its usefulness as the primary measure of students. There are many reasons for undesirability of traditional (or teacher) assessment in which student' knowledge is evaluated by one or two single scores. This element makes students rely on their memorization ability and reproduce these pieces of information from their memory on the exam to score high and after the exam this information disappeared. This traditional assessment distracts the students from meaningful learning. Also many other factors, among other things, distraction, anxiety and stress may influence students' performances.

Recent approaches towards assessment stress the learning potential of assessment (Taras, 2008). This is labeled as formative assessment and defined as "assessment that is specifically intended to provide feedback on performance to improve and accelerate learning" (Nicol & Milligan, 2006, p. 64). Some consider this as a key quality of assessment and regard this as the "consequential validity" of assessment (Gielen et al., 2003). Consequential validity is put next to the two other traditional psychometric qualities of an assessment: reliability and validity. According to Messick (1994) consequential validity is one of the six aspects of his unified concept of validity.

Today, there are innovations in assessment procedures, where the change is from summative assessment to formative assessment. These innovations involve thinking of alternatives, which require questioning the learning process and using learning and assessment activities together rather than habitual testing applications. Alternative assessment asks students to show what they can do, that is to say, students are evaluated on what they integrate and produce rather than on what they are able to recall (Coombe et al., 2007). A large number of novel approaches of assessment have hence been suggested which meant to develop the integration of learning and assessment by enhancing the engagement of students in the assessment tasks (Sluijsmans et al., 2003).

REVIEW OF LITERATURE

The literature suggests that students need to develop as independent learners in order to be successful in their higher education programs and also in their professional lives post-graduation. Boud and Falchikov (2007) have described the ability to evaluate one's learning and performance as an essential part of "becoming an accomplished and effective professional" (p. 184). Similarly, Biggs and Tang (2007) argued that the ability to make judgments about whether a performance or product meets a given criteria is vital for effective professional action in any field. Tan (2007) also argued for "self-assessment development practices that can develop and sustain students' self-assessment ability beyond its immediate programme of study" (p. 115). However, part of this preparation for the future requires helping students to learn to continuously monitor the quality of their work during the act of production itself, so they can make improvements in real time (Montgomery, 2000). Two effective teaching and learning processes that can assist with the development of such judgment are self-assessment and peer-assessment, and the literature provides examples of how these processes have been used successfully in education.

Involvement of students in assessment can be organized in two ways: peer and self-assessment. 'Peer-assessment', as one of the main forms of alternative assessment, has gained much attention in educational learning and educational research. It is considered as an arrangement in which individuals consider the amount, level, value, worth, quality, or success of the products or outcomes of learning of peers of similar status (Topping, 1998). It is "the process of having the readers critically reflect upon, and perhaps suggest grades for the learning of their peers" (Roberts, 2006, p. 80), and being judged for the quality of the appraisals made (Davies, 2006). Immediate support in the classroom, gains for both the assessor and the assessed, and being individualized and interactive are some benefits of peer-assessment to consider (Black & William, 1998).

Saito (2008) believes peer-assessment encourages reflective learning through observing others' performances and becoming aware of performance criteria. In general, peer-assessment seems to generate positive reactions in students, although some students have concerns and worries, it leads to the development of self-awareness, noticing the gap between one's and others' perception, and facilitating further learning and responsibility for it. In addition, focusing on peers' strengths and weaknesses can enhance students' learning, raise their level of critical thinking, and lead them to autonomy. According to Zhi-Feng and Yi Lee (2013), the students made positive modifications to their work with the help of feedback from others after participating in peer-assessment activities. Most of the students had positive opinions regarding peer observation.

Based on the new developments in learning theories, teachers open up discussion of assessment with students; this is actually what presents a major challenge for assessment in 21st century because it is putting demands on the teacher to obtain

specific skills needed for this new, additional role. The process of learning should be assessed by more intense, interactive methods and that work should be undertaken in collaboration, either between teacher and student or a group of peers (Matsuno, 2009; Wikstorm, 2007).

Boud (1995) stresses that the assessment process shouldn't be thought only as an instrument to give students a diploma, but it should also be a process that leads up to student development and better learning conditions and applications. Such alternative views on assessment have given rise to new approaches like self-assessment. It has been argued that 'self-assessment' serves as an effective language learning strategy to promote autonomous language learning because it encourages language learners to assess their learning progress and in turn helps them to stay focused on their own learning (Chamot & O'Malley, 1994; Chen, 2005; O'Malley & Pierce, 1996; Oscarson, 1997). The proponents of self-assessment strategies maintain that participating in self-assessment can help learners become skilled judges of their own strengths and weaknesses and establish realistic and attainable goals for themselves, thus developing their self-directed language learning ability (Chamot & O'Malley, 1994; Dickinson, 1987; Oscarson, 1997).

Tavakoli (2010) argued that in self-assessment, students concurrently create and undergo the evaluation process, judging their achievement in relation to themselves against their own personal criteria based on their own objectives and learning expectations. Matsuno (2009) is of the belief that self-assessment can give students a chance to build up their experiences in language learning and this experience can motivate students to be more involved in the classroom because they feel that they have control in their own learning rather than just having the teachers tell them what they have to learn. It also provides an opportunity for English (as a foreign or second language) learners to monitor their own progress and take responsibility for meeting goals. Therefore, self-assessment brings autonomy for learners. Portfolio assessment that is one type of self-assessment also fosters learners' autonomy that may contribute to enhancing motivation and language learning (Hosseini & Ghabanchi, 2014). Weisi and Karimi (2013) found a significant effect of self-assessment initiatives in enhancing the students' willingness and ability to engage in self-assessment and in creating positive outlooks toward English language learning.

To conclude, peer and self-assessment are the alternatives in language assessment. In peer-assessment, according to Falchikov (2005, p. 27), "students use criteria and apply standards to the work of their peers in order to judge that work". Building on the latter, in self-assessment students use criteria and apply standards to judge their own work. Both peer and self-assessment are expected to decrease the central role of the teacher in assessment activities. During the last decades, there has been an increase in the implementation of peer and self-assessment in higher education learning environments (Cheng & Warren, 2005; Glyn et al., 2011; Matsuno, 2009; Patri, 2002; Tavakoli, 2010; Weisi & Karimi, 2013; Wikstorm, 2007).

In line with previous studies, although not aiming at reviewing and replicating the extensive literature on peer and self-assessment, the present study is conducted to shed light on the status of peer, self, and traditional assessment in Iranian classrooms where teacher-centered classes are the norm. Considering the importance of reading comprehension in an EFL context, the current study focuses on investigating the effect of these three types of assessment on Iranian learners' L2 reading comprehension.

Even though there are several ways for assessing reading comprehension in an EFL context, using peer and self-assessment for L2 text comprehension has not been appreciated at least in Iranian EFL classrooms. As a result, the need arises to study the effect of peer, self, and traditional assessment on the students' achievement in comprehension of L2 texts. Some researchers report that there are several shortcomings and limitations among traditional testing methods. Traditional assessment involves the employment of paper-and-pencil tests and standardized tests to assess student's performance under time pressure. Typically, traditional or teacher assessment is used only to monitor students' learning. Under this model, students who know are separated by those who do not know. In other words, traditional assessment creates a system that classifies and ranks students (Berlak, 1992, Stiggins, 2001).

In traditional assessment, generally the teacher alone has the power to make decisions about what is learned and how it is assessed and students do not participate in making decisions about what is important for them to learn or in determining how well they are learning (Heron, 1988; Sessions, 1995). But, the focus of alternative assessment is on developing real-world problem solving skills that will lead people to observe, think, question, and test their ideas (Herman et al., 1992). Alternative assessment embraces a democratic decision-making process (Heron, 1988). In contexts that use alternative assessment practices, students and instructors are co-learners, freely expressing and testing their ideas together.

There are many kinds of alternative assessment like peer-assessment, self-assessment, play-based assessment, conference assessment, and so on. The educational system in Iran is based on traditional or teacher assessment and rote learning. This traditional assessment is not authentic and does not demonstrate actual level of proficiency. In this study, peer, self, and traditional assessment were selected as tools and the skill be assessed is L2 reading comprehension. Reading is one of the four major skills in learning a foreign language and the one that provides the students with the best opportunity of being in contact with English after education. The teacher researcher's presupposition in this study is that students' difficulties in reading comprehension can be at least minimized if she uses peer and self-assessment for assessing of EFL learners because this way facilitates the learning process, enhances peer and self-directed learning, encourages learner's autonomy, raises learner's awareness about learning strategies, and improves learners' reading comprehensions ability. Therefore, the researchers try to investigate the effect of peer, self, and traditional assessment on the

Iranian learners' L2 reading comprehension in order to have an empirical evidence of such an effect.

THIS STUDY

The purpose of this study is threefold: first, it attempts to investigate the impact of peer, self, and traditional assessment on EFL learners' reading comprehension ability; second, it aims to enable teachers and students to share the responsibility for setting learning goals and for evaluating progress toward meeting those goals; third, it may help students become peer and self-directed learners; teachers are no longer knowledge transmitters but mentors, facilitators and collaborators. Students can become active learners by taking more responsibility in learning and having more involvement in assessment. It may also help students to become realistic judges of their own performance, by enabling them to monitor their peer and own learning, rather than relying on their teachers for feedback. The goal is to compare peer, self, and traditional assessments with one another and decide which is more suitable and effective for the students in promoting L2 reading comprehension in an Iranian EFL classroom setting. According to the stated problem and the purpose of the study, the following research question is addressed:

Is there any significant difference between the three modes of assessment such as peer, self, and traditional in terms of their effects on EFL learners' L2 reading comprehension ability?

METHOD

The present study was carried out in an EFL classroom. A quantitative research design was employed due to the nature of this research and the research question.

Participants

The population from which the participants were selected was 102 female EFL learners who were aged between 17-21 years old in a private English Language Institution in Isfahan. To assess their general language proficiency level, the standard test of Nelson (2001) was administered. The students' performance on the reading comprehension section of Nelson test was analyzed to ensure that they were homogeneous in terms of their proficiency level. Only the participants whose scores on this test fell between one standard deviation above and one standard deviation below the mean was selected as the sample of the study. Finally, 77 participants were qualified to be included in this study. Later, these homogenized participants were randomly assigned to three experimental groups who subsequently exposed to peer, self, and traditional assessment.

Materials

Nelson test

In order to determine the general proficiency level of the participants and to screen them, reading section of Nelson test (2001) including 30 questions, was used for the selection of 77 intermediate participants. The individual scores on this section of Nelson were analyzed to ensure that they were of the same level of language proficiency. These three groups were almost equal regarding reading comprehension ability at the beginning of the study.

Achievement comprehension test

A 24 item multiple-choice test of reading comprehension following 4 reading passages (including the same title and the same key words and expressions but different in content from reading passages of treatment) was used as an achievement comprehension test for experimental groups to find out the learners' reading comprehension ability at the end of the treatment. The reliability and validity of the reading comprehension test were established. Four specialists in language teaching and testing were asked to review the test, and there was a general consensus among them concerning the content and face validity of the test. In order to ensure that the achievement test was reliable, KR-21 reliability method was used in this study, and it was 0.86.

Procedure

After the teacher researcher made certain that the participants form a homogenous sample, they were randomly assigned to three experimental groups. On the day of the exposure to the treatment, the experimental groups received peer, self, and traditional assessment on L2 reading comprehension. In the first meeting with the participants of experimental groups, the teacher researcher presented the idea of peer, self, and traditional (or teacher) assessment, the purpose, the basic components and the procedures of these assessments for each experimental group respectively. All groups were asked to read 4 reading passages in the class and answer its comprehension questions. Then, the peer, self, and traditional assessment process was subsequently done in each experimental group.

Reading Logs for each passage were used as peer, self-assessment tool for monitoring the reading comprehension and strategy use, questing students' progress over time, evaluating the reading passages, reflections about the various reading challenges students (peers or individuals) faced, the different approaches they experimented with, and summarizing the whole text into an appropriate reading strategy chart.

In order to depict the three groups' performance and to examine the effect of treatment (peer, self, and traditional assessment) on L2 reading comprehension, in the next

session which was held seven days after treatment, the 24 item multiple-choice test of reading comprehension following 4 reading passages was used as an achievement comprehension test for each experimental group. It should be mentioned that the contents of reading passages of the achievement comprehension test were different from those of the treatment; however, they include the same title and the same key words and expressions.

Data analysis

The raw scores of the 77 participants were compiled for data analyses. Descriptive statistics were used to determine the mean and standard deviation of each group on the achievement reading comprehension test. One-way analysis of variance (ANOVA) was performed to compare the groups on the basis of outcome measures at the .05 level of significance. ANOVA accomplishes its statistical testing by comparing variance between the groups to the variance within each group. A significant statistical finding would indicate that group means were significantly different from each other. In case of a significant statistical finding, there is a need to use a Post-Hoc test (Tukey, Scheffe, Bonferroni or others) to find exactly which groups differed from which other groups (Balian, 1994). In this research, because of a significant finding from ANOVA, Bonferroni test was used to find exactly which groups differed from each other. In other words, one-way ANOVA was employed to calculate the amount of variance between and within the three groups, and Post hoc test was run to determine whether the difference existing among groups was significantly meaningful for peer, self, and traditional assessment.

RESULTS

The research question concerned the effect of peer, self, and traditional assessment on improving EFL learners' reading comprehension. Table 1 represents the descriptive statistics and ANOVA results of the achievement reading comprehension test for three experimental groups.

Table 1. One-way ANOVA of the comprehension test for the three groups

Groups	N	Mean	SD		df	F	Sig.
Peer-assessment	25	6.88	2.00	Between groups	2	5.475	.006
Self-assessment	26	5.46	2.16	Within groups	74		
Traditional assessment	26	4.81	2.62	Total	76		

According to Table 1, since the significance level (.006) is smaller than the alpha level (.05), there were significant differences among the three groups (peer, self, and traditional). As the results suggest, types of assessment in this study appear to have a differential effect on the learners' abilities to comprehend L2 texts. The Post hoc Bonferroni test in the following table shows the level of significant differences among three groups.

Table 2. Post-Hoc Bonferroni test results for comprehension scores

(I) condition	(J) condition	Mean Difference (I-J)	Std. Error	Sig.
Peer-assessment	Self-assessment	1.4185	.6390	.089
	Traditional	2.0723	.6390	.005*
Self-assessment	Peer-assessment	-1.4185	.6390	.089
	Traditional	.6538	.6327	.914
Traditional assessment	Peer-assessment	-2.0723	.6390	.005*
	Self-assessment	-.6538	.6327	.914

As shown in Table 2, the level of significance for peer-assessment group and self-assessment group is .089, meaning that no significant difference exists between these two groups in reading comprehension test. The level of significance for peer-assessment group and traditional assessment group is .005, which means there is a meaningful difference in the performance of the groups on peer-assessment comprehension measures. Based on the results of the post hoc test, the level of significance for self-assessment group and traditional assessment group is $p > .05$. Thus, this conclusion can be drawn that there exists no significant difference between these two groups. The results indicate that a meaningful difference was found only between the two assessment types (peer and traditional) for the reading comprehension test.

DISCUSSION

With respect to the effect of peer, self, and traditional (or teacher) assessment on improving EFL learners' reading comprehension, the results of the present study show a significant statistical difference in reading comprehension between the performance of peer-assessment group and traditional assessment group which means that peer-assessment results in better reading comprehension for Iranian EFL learners. This finding supports the findings of (Langan et al., 2008) that peer-assessment enhances learning process more than traditional assessment. According to Matsuno (2009), the merits attributed to applying peer-assessment cannot be ignored. Peer-assessment is an effective tool in both group and individual projects, and can encourage reflective learning through observing others' performances and awareness of performance criteria, but traditional or teacher assessment cannot (Ballantyne et al., 2002; Saito, 2008).

Regarding the impact of peer-assessment and traditional assessment on L2 text comprehension, the finding of this study confirm the previous findings (Patri, 2002; Saito, 2008; Yamashiro, 1999) that learners could rate their peers acceptably and improve their judgments so that they could acquire a better understanding of their own skills. According to Langan et al., (2008) the lower intra-class correlation suggested that peers and teachers still interpret the criteria and indicators of the rubric in a different way. Within the group of peers, not all students could have applied the

same criteria in a consistent way. Lastly, the peers report higher marks as compared to teachers. The finding of the current study also is in agreement with the results of other study (Praver et al., 2011) in that peer-assessment wherein students provide comments are regarded, both in their production and reception, as more useful for English reading comprehension skill development.

Concerning the impact of self-assessment and traditional (or teacher) assessment on L2 text comprehension, the results of this study indicated that the self-assessment group outperformed the traditional assessment group but not significantly. This finding is in line with the study conducted by Patri (2002). He found that the self-assessment scores are, mostly, higher than the marks given by teachers in traditional assessment. The findings of the present study is in consistent with the results of other studies which report lower correlations values between self and traditional (or teacher) assessments as compared to the correlation values between traditional and peer-assessment (Campbell et al, 2001; Langan et al., 2008; Patri, 2002). Nevertheless, others consider self-assessments to be as valid as peer-assessment (AlFallay, 2004; Hafner and Hafner, 2003). Tavakoli (2010) believes that self-assessment would mitigate the student teacher relationship by giving responsibility to students as to their own progress and to their own learning would so that they can become more motivated in participating in their evaluation for future learning expectations. It has been argued that self-assessment serves as an effective language learning strategy to promote autonomous language learning because it encourages language learners to assess their learning progress and in turn helps them to stay focused on their own learning (Chamot & O'Malley, 1994; Chen, 2005; O'Malley & Pierce, 1996; Oscarson, 1997).

With regard to the comparison of self-assessment scores and peer-assessment scores, there was no meaningful difference between the performance of self-assessment and peer-assessment groups in L2 text comprehension measures. As the results of the present study showed the peer-assessment group insignificantly outperformed the self-assessment group in L2 reading comprehension. This finding is in agreement with the findings of (Segers et al., 2003) that both self and peer-assessment are expected to decrease the central role of the teacher in assessment activities. Topping (2003) additionally mentions economic benefits to adopt self and peer-assessment. Shifting part of the responsibilities for assessment and feedback from the teacher to the student has, next to educational benefits, also benefits in terms of reducing teaching workload.

Weisi and Karimi (2013) found that new and alternative forms of assessment such as self or peer-assessment can be beneficial in language learning. Self and peer-assessment result in a more active involvement of students in their own learning process (Ozogul & Sullivan, 2009). A student who always expects teachers to present a judgment will develop, to a lesser extent, a self-assessment orientation (Boud & Falchikov, 2007). Topping (2009) explains this by linking peer-assessment to the provision of immediate, individualized and richer feedback. Since this feedback is formative in nature, it has a clear potential of fostering the subsequent learning process (Hattie, 2009).

CONCLUSION

This research study made attempts to launch an investigation into the effects of peer, self, and traditional (or teacher) assessment on improving L2 learners' reading comprehension ability. Three main conclusions can be drawn from the findings. Firstly, there was statistically a significant difference between the performances of the peer and traditional assessment groups in terms of their reading comprehensions ability. The comparisons made showed that the reading comprehension of those students in the group where peer-assessment implemented differed significantly from those of students in the group where traditional or teacher assessment implemented. Secondly, there was no significant difference between the impact of self-assessment and traditional assessment on EFL learners' reading comprehension ability. Thirdly, no meaningful difference was also found in the performance of self-assessment and peer-assessment groups in L2 reading comprehension.

Peer assessment recently has been regularly practiced as an alternative assessment technique, predominantly in higher education. Peer-assessment is an integral segment of learning experiences. During the course of the assessment, students learn to shoulder high levels of responsibility and to concentrate on learning itself. Peer assessment also offers the learners a situation in which they can perceive the role of their teachers and appreciate the nature of assessment. It supports students to learn about learning and, consequently, increases students' meta-cognitive understanding about their own learning.

The results of the present study have several implications. This study helps students become peer and self-directed, and enables teachers to be facilitators and collaborators. Students are able to witness their peers and arrive at a better understanding of how their peers learn. They are autonomous learners. They take more responsibility in their own learning, and have more involvement in assessment. Syllabus designers can get the insight from the results of this study in designing an appropriate syllabus. Syllabus designers should also consider and value learners' right for their own decisions and suggestions and criticism while designing syllabuses. In the current study, the findings promote the meaningful learning and decrease the problems of the rote learning. They also showed that peer-assessment is more useful in teaching and meaningful learning.

There have been some limitations in conducting this study. In the present study, the researcher was in the shortage of time, and the study lasted in a few weeks. This study was conducted only with 77 participants in a private language institute in Isfahan, and the researcher also had to select only from female students with average age of 17-21 years old. Similar studies can change variables of age and gender. Finally, among different types of alternative assessment, just peer and self-assessment, and among all skills of language learning, just reading, have been taken into account in this study. So, interested researchers can investigate the impact of other types of alternative

assessment techniques such as portfolio and conference assessment on other language skills such as listening, speaking and writing.

REFERENCES

AlFallay, I. (2004). 'The role of some selected psychological and personality traits of the rater in the accuracy of self and peer-assessment'. *System, 32,* 407-425.

Balian, E. S. (1994). *The graduate research guidebook: A practical approach to doctoral/masters research.* Lanham, MD: University Press of America.

Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer-assessment in large classes using an action research process. *Assessment and Evaluation in Higher Education, 27,* 427-441.

Berlak, H. (1992). The Need for a new science of assessment. In H. Berlak and others (eds.), *Toward a new science of educational testing and assessment.* Albany: State University of New York. Retrieved from books.google.com.

Biggs, J., & Tang, C. (2007). *Teaching for quality learning at university: What the student does* (3rd ed.). Maidenhead, Berkshire: Open University Press.

Black, P. & William, D. (1998). *Inside the black box: Raising standards through classroom assessment.* London: Department of Education and Professional Studies, Kings College.

Boud, D. (1995). *Enhancing learning through self-assessment.* Kogan Page, London.

Boud, D., & Associates. (2010). *Assessment 2020: Seven propositions for assessment reform in higher education. Sydney*: Australian Learning and Teaching Council.

Boud, D., & Falchikov, N. (2007). Developing assessment for informing judgment. In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education: Learning for the longer term* (pp. 181-197). London: Routledge.

Campbell, K., Mothersbaugh, D., Brammer, C., & Taylor, T. (2001). Peer versus self-assessment of oral business presentation performance, *Business Communication Quarterly, 64*(3), 23-42.

Chamot, A. U., & O'Malley, J. M. (1994). The CALLA Handbook: *Implementing the cognitive language learning approach reading*, MA: Addison Wesley. Retrieved fromjillrobbins.com/calla/references .html.

Chen, Y. M. (2005). *Peer and self-assessment in EFL classrooms.* Selected Papers of the Fourteenth International Symposium on English Teaching (pp. 320-329). Taipei: Crane Publishing.

Cheng, W, & Warren, M. (2005). Peer-assessment of oral proficiency. *Language Testing, 22,* 93–121.

Collins, J., & O'Brien, N. (2003). *The Greenwood dictionary of educatio*n. USA, Westport: Greenwood Press. September, 2001. Retrieved February 24, 2007, from: http://www.leeds.ac.uk/educol/ documents/ 00001862.htm.

Coombe, C, Folse, K, & Hubly, N. (2007). *Assessing English language learners*. United States of America: University of Michigan Press.

Crooks, T. (2001). *The validity of formative assessment*. Paper presented to the British educational research association annual conference, University of Leeds, 13-15.

Davies, P. (2006). *Peer-assessment*: judging the quality of students' work by comments rather than marks. *Innovations in Education and Teaching International*, *43,* 69–82.

Dickinson, L. (1987). *Self-instruction in language learning*. Cambridge: Cambridge University Press.

Falchikov, N. (1995). Peer feedback marking: Developing peer-assessment. *Innovations in Education and Training International, 32,* 175–187.

Falchikov, N. (2005). *Improving assessment through student involvement. Practical solutions for aiding learning in higher and further education*. New York: Routledge Falmer.

Gielen, S., Dochy, F., & Dierick, S. (2003). Evaluating the consequential validity of new modes of assessment: the influence of assessment on learning, including pre-, post-, and true assessment effects, in M. Segers, F. Dochy, & E. Cacallar (Eds). *Optimising new modes of assessment: In search of qualities and standards*, (pp. 37-54). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Glyn, T; Dona, M; & Kathleen, P. (2011). Using self- and peer-assessment to enhance students' future-learning in higher education., *J ournal of University Teaching & Learning Practice, 8*(1).

Gronlund, N., & linn, R. (1990). *Measurement and evaluation in teaching*. NewYork: Macmillan.

Hafner, J., & Hafner, P. (2003) Quantitative analysis of the rubric as an assessment tool: an empirical study of peer-group rating, *International Journal of Science Education*, *25*(12), 1509-1528.

Hattie, J. (2009). *Visible Learning: A Synthesis of over 800 Meta-analysis relating to Achievement*. Milton Park, Oxon: Routledge.

Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment*. Retrieved fromwww.e-lead.org/principles/structure.asp.

Heron, J. (1992). Assessment Revisited. In D. Boud (ed.), *Developing student autonomy in learning* (2nd ed.). York: Nichols.

Hosseini H., and Ghabanchi Z. (2014). The Effect of portfolio assessment on EFL learners' reading comprehension and motivation. *English Language Teaching, 7*(5).

Langan, M., Shuker, D., Cullen, R., Penney, D., Preziosi, R., & Wheater, P. (2008). Relationships between student characteristics and self and peer and tutor evaluations of oral presentations, *Assessment & Evaluation in Higher Education, 33*(2), 179-190.

Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing, 26*, 75–100.

Messick, S. (1994). *Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning*. Research Report RR 94-95. Princeton: Educational Testing Service.

Montgomery, K. (2000). Classroom rubrics: Systematizing what teachers do naturally. *Clearing House, 73*(6), 324-328.

Nelson, T. (2001). *Practice Test II.* Hong Kong: Bright Sun Printing Press.

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.

Nicol, D., & Milligan, C. (2006). Rethinking technology-supported assessment practices in relation to the seven principles of good feedback practice, in C. Bryan & K. Clegg (Eds.). *Innovative assessment in higher education*. (pp. 64-77). Taylor and Francis Group, London.

O'Malley, M. J., & Valdes Pierce, L. (1996). *Authentic assessment for English language learners*: *Practical approaches for teachers*. New York: Addison Wesley Publishing Co.

Oscarsson, M. (1997). Self-assessment of foreign and second language proficiency. In C. Clapham and D. Corson (eds), *Encyclopedia of Language and Education, (V.7): Language Testing and Assessment* (pp. 175-187). Netherlands: Kluwer Academic Publishers.

Ozogul, G., & Sullivan, H. (2009). Student performance and attitudes under formative evaluation by teacher, self, and peer evaluators. *Educational Technology Research and Development, 57*, 393-410.

Patri, M. (2002). The influence of peer feedback on self and peer-assessment of oral skills. *Language Testing, 19,* 109–131.

Praver, M., Rouault, G., & Eidswick, J. (2011). Attitudes and affect toward peer evaluation in EFL reading circles. *The Reading Matrix. 11*(2), 89-101.

Roberts, T. (2006). *Self-, peer-, and group assessment in E-learning.* United States of America: Information science publishing.

Saito, H. (2008). EFL classroom peer-assessment: Training effects on rating and commenting. *Language Testing, 25,* 553–581.

Segers, M., Dochy, F., & Cascallar, E. (2003). The era of assessment engineering: changing perspectives on teaching and learning and the role of new modes of assessment. In M. Segers, F. Dochy, & E. Cascallar (Eds). *Optimising new modes of assessment: In search of qualities and standards,* (pp. 1-12). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Sessions, R. (1995). *Education is a gift, not a commodity.* Paper presented at the National Conference of the Community Colleges Humanities Association, Washington, D. C. Retrieved from: www.jcu.edu/academic/Different%20ways%20to%20Grade.pdf

Sluijsmans, D. M. A., Brand-Gruwel, S., vanMerrienboer, J. J. G., & Bastiens, T. J. (2003). The training of peer-assessment skills to promote the development of reflection skills in teacher education. *Studies in Educational Evaluation, 29,* 23-42.

Stiggins, R. (2001). *Student- involved classroom assessment.* N.J., Upper Saddle River Prentice-Hall.

Tan, K. (2007). Conceptions of self-assessment: What is needed for long term learning? In D. Boud & N. Falchikov (Eds*.), Rethinking assessment in higher education: Learning for the longer term* (pp. 114-127). London: Routledge.

Taras, M. (2008). Summative and formative assessment: perceptions and realities, *Active Learning in Higher Education, 9*(2): 172-192.

Tavakoli, M. (2010). Investigating the relationship between self-assessment and teacher-assessment in academic contexts: A Case of Iranian university students. *The Asian EFL Journal Quarterly, 12*(1), 234-260.

Topping, K. J. (1998). Peer-assessment between students in colleges and universities. *Review of Educational Research, 68,* 249–276.

Topping, K. (2003). 'Self- and peer-assessment in school and university: reliability, validity and utility', In M. Segers, F. Dochy, & E. Cacallar (Eds), Optimising new

modes of assessment: In search of qualities and standards, (pp. 55-87). Dordrecht, The Netherlands: Kluwer academic publishers.

Topping, K. (2009). Peer-assessment, *Theory into Practice, 48,* 20-27.

Weisi, H., Karimi, M. (2013). The effect of self-assessment among Iranian EFL learners. *Procedia - Social and Behavioral Sciences, 70*(25), 731–737.

Wikstorm, N. (2007). *Alternative assessment in primary years of international baccalaureate education* (Master's thesis).

Yamashiro, A. (1999). *Using structural equation modeling to validate a rating scale.* Presented at the 21st Language Testing Research Colloquium, Tsukuba, Japan.

Zhi-Feng, E., & Yi LEE, C. (2013). Using peer feedback to improve learning via online peer-assessment. *Turkish Online Journal of Educational Technology, 12*(1), 187-199.