

On the comparative impact of self-assessment and peer assessment on Iranian male and female EFL learners' accuracy in speech

Abdollah Moghimi¹, Islamic Azad University, Department of Humanities, Larestan
Branch, Iran

Suggested Citation:

Moghimi, A. (2022). On the comparative impact of self-assessment and peer assessment on Iranian male and female EFL learners' accuracy in speech. *Contemporary Educational Researches Journal*. 12(4), 204–213. <https://doi.org/10.18844/cerj.v12i4.7468>

Received from March 12, 2022; revised from June 25, 2022; accepted from August 17, 2022.

Selection and peer-review under the responsibility of Assoc. Prof. Dr. Deniz Ozcan, Ondokuz Mayıs University, Turkey.
©2022 Birlesik Dunya Yenilik Arastirma ve Yayıncılık Merkezi, Lefkosa, Cyprus.

Abstract

The current study tried to explore the effects of assessment types and learners' gender on EFL learners' performance in speech accuracy. To this end, 60 homogenous students out of 90 students from Rahe Farda Institute in Bastak in Hormozgan province were selected based on the Quick Oxford Placement Test. The instruments of the study were an Oxford Placement Test, peer and self-assessment questionnaires adopted from Patri and Weir's Analytic Speaking Criteria. The results were calculated using SPSS version 20. It can be seen that the means were close together, but the mean score for the male students was a little higher than that of the female students. According to the outcomes, assessment types had a significant influence on performance in speech accuracy and peer-assessment was more effective than self-assessment on performance in speech accuracy.

Keywords: Peer assessment, self-assessment, speech;

* ADDRESS FOR CORRESPONDENCE: Abdollah Moghimi, Islamic Azad university of Larestan, Lar, Iran.
E-mail address: abdollahmoghimi85@gmail.com

1. Introduction

It is generally claimed that a large number of assessment strategies are essential for evaluating the learners' performances (Birjandi & Hadidi Tamjid, 2012). In addition, students should be aware of their abilities and they should know whether they can use the skills they have acquired (Javaherbakhsh, 2012; Gunduz et al, 2017). Being unaware of this capacity makes them inefficient in learning. It is also stated that from an educational point of view, understanding the students' performances would be important for them, and it makes the teachers aware of their individual needs. In the Iranian educational system, learning English is started from junior high school and it is continued until the end of the pre-university courses. In other words, the students have the opportunity to study English for years in formal education systems.

In addition, it seems that those learners take part in private language centres and spend more time and energy on learning English. However, in spite of the large amount of time and budget assigned to learning English, most of the students have difficulty in this subject area (Hooshangi, Yousofi, & Mahmoudi, 2014). In the area of oral proficiency skills, the problem seems more severe and the Iranian learners of a foreign language have more difficulty. The students' performances are still assessed by the teachers at schools, and the students have a few chances to assess their own performances. In this regard, Luoma and Tarnanen (2003) refer to self-assessment as a tool that helps students understand the aims of the tasks, 'reflect on what they have achieved with reference to such goals, and figure out what it will take to finally reach their goals'. In spite of the significance of self-assessment, students do not often rate their own performances (Housen & Kuikken, 2009).

According to Blue (1994, as cited in Coombe & Canning, 2016), interest in self-assessment was obtained as a result of a general interest in the area of autonomous learning. Self-assessment has received great attention from assessing the learners' language competencies. In addition, self-assessment can monitor the learners' performance on a language learning task 'after implementation or checking their success in using a language' (Richards & Schmidt, 2002). They also believed that self-assessment can be an example of a meta-cognitive strategy in second language learning. Self-assessment is an assessment technique that refers to the process, whereby 'learners simultaneously create and undergo the evaluation procedure, judging their achievement in relation to themselves against their own personal criteria, in accordance with their own objectives and learning expectations' (AbolfazliKhonbi & Sadeghi, 2012).

Assessment is an issue that influences different aspects of individuals' lives (Kargar, Salimi, & Zareian, 2014). It also affects the lives of those people dealing with children and those who work at schools. Assessment plays a key role in both learning and teaching. It also acts as a motivator for the learners and a tool for teachers (Lindsay & Knight, 2006). Since assessment may affect learning, it is considered an inevitable factor in all educational systems (AbolfazliKhonbi & Sadeghi, 2012). In addition, assessment may influence motivation through meaningful involvement of students in the learning process. Furthermore, through the assessment process, teachers are able to recognise the students' weaknesses and strengths. Assessments can also be made valid, fair, ethical, feasible and efficient tools for learning using multiple measures (Mousavi, 2012).

In line with the present study, Liang et al. (2020) evaluated the evaluation of nurses' professional competence and compared the evaluations using several methods: self-evaluation, peer evaluation and supervisor evaluation. Nurse practitioners (NPs) require an increased effort to continuous learning for enhancing professional competencies. Moreover, using multiple methods for assessment to obtain a more comprehensive and accurate evaluation of NPs' professional competence (Liang, Tang, Wang, & Yu, 2021). Lee et al. (2021) studied the reliability, effectiveness and benefits of peer evaluation in Chinese college English language classes. The results indicated that the reliability of up to two peer assessors' holistic scoring is equivalent to that of one teacher assessor's scoring; furthermore, qualitative comments made by peer and teacher assessors were similarly effective; finally, peer assessment is beneficial for assessment receivers

and peer assessors. Implications for college EFL teachers and students are discussed (Li, Huang, & Cheng, 2022). Colbran (2017) studied the evaluation of the usefulness of self-assessment, peer assessment and academic feedback mechanisms (Colbran, 2017).

As for speaking ability, Ahangari, Rassekh-Alqol and Ali Akbari (2013) examined the effect of peer assessment on oral presentation of Iranian EFL students. The results of their study specified a statistically significant difference among the groups. In addition, the findings of their study showed that peer assessment can enhance the speaking ability of learners if they are given a valid criterion or framework (Ahangari et al., 2013). Meihami and Varmaghani (2013) also investigated the use of self-assessment in EFL writing classroom. The results of their analysis showed that self-assessment significantly improved the writing proficiency of the participants. Their results confirmed that self-assessment is successful with EFL students (Meihami & Varmaghani, 2013). In another study, Ariaifar and Fatemipour (2013) examined the effect of self-assessment on the speaking skill of pre-intermediate Iranian EFL learners. The results of their study indicated that self-assessment practices helped the participants of the study to improve their speaking ability (Ariaifar & Fatemipour, 2013).

According to Javaherbakhsh (2010), 'along with these problems of learning, the practices of self- and peer-assessment also reflect new thinking about classroom assessment: assessment for learning and evaluation' (Javaherbakhsh, 2012). It seems that the focus has been changed from summative to formative assessment to prove that the students have this ability to provide feedback and it reveals that they are making progress (Chen, 2008). Generally, syllabus and course designers, curriculum developers and university instructors in particular should pay more attention to students' needs and styles, since they may have crucial roles as contributing factors to the students' ultimate academic success (Chen, 2008). Hence, conducting a research is necessary to let the students gain autonomy in and self-awareness of their learning. As a conclusion, this study attempts to investigate whether learners' self- and peer-assessment of the learning process influence the development of their accuracy in speech.

2. Methods

Participants in this study were male and female students of Rah-e-Farda Institute in Bastkin province of Hormozgan in the age range of 19–25 years. There were 90 learners who were given an OPT (Oxford Placement Test) homogenisation test. Participants whose scores ranged from a standard deviation above and below average were selected. About 60 students (36 girls and 24 boys) were selected from among 90 students as participants in the present study. Sampling method of this research is available sampling or available. Oxford Level Test Questionnaire, Peer Assessment Questionnaire, Self-Assessment Questionnaire, Weir Analytical Speech Criteria and Pre-Speech Strategies (Cohen, 1996) were used to evaluate participants and to analyse data obtained from descriptive and inferential tests (analysis of variance). An independent sample *t*-test was carried out using SPSS software version 20. Prior to the main study, a Quick Oxford Placement Test (QPT) was used to homogenise the participants of the study. Therefore, all initial 90 students took part in QPT and students whose scores were between one standard deviation below and one standard deviation above the mean participated in the main study. Table 1 shows the descriptive statistics of the participants related to QPT. In the present study, ethical considerations have been observed and participants have been assured that their information is safe.

3. Instruments

3.1. Oxford placement test

To identify the proficiency level of the students participating in the study, and to be sure whether the students were homogeneous, an Oxford placement test was used by the researcher. This placement test is a standardised validated and reliable placement test which uses inter-rater reliability, being used by prestigious schools and universities all around the world. Each test includes two sections (listening and

grammar). 'These sections are also integrated with reading skills and vocabulary in context' (Van Loon, 2007). According to Allen (2004), the QPT is a valid and reliable test and a highly effective instrument in grouping students into appropriate levels. It can also be used as a quick measurement of students 'general language proficiency'. It includes two parts: part one (questions 1–40) and part two (question 41–60). According to Geranpayeh (2003), the test has been validated in 20 countries by more than 60,000 students. There were three validation phases in the course of QPT development. In addition, Geranpayeh (2003) reports that the typical reliabilities of the QPT have been calculated during the three validation phases; they are close 0.9 for the 60-item test and 0.85 for the 40-item test.

3.2. Peer-assessment questionnaire

This questionnaire was adopted from Patri (2002). It is a Likert-type questionnaire with five options from 1 (poor) to 5 (excellent). The reliability of the instrument was reported to be 0.70 (Saito, 2008). The questionnaire was validated by Sluijsmans (2002). The questionnaire included 14 questions in 4 categories. Category 1 comprised questions 1–6 that composed of organisation and content of presentation. Category 2 (questions 7–9) was about the use of language. Category 3 (questions 10–12) was related to manner. Finally, Category 4 contained questions 13 and 14 which were related to the interaction with the audience (Sluijsmans, 2002).

3.3. Self-assessment questionnaire

The same questionnaire was used for self-assessment with slight adjustments. As mentioned earlier, the instrument was adopted from Patri (2002). The reliability of the instrument was reported to be 0.70 (Saito, 2008), and was validated by Sluijsmans (2002). This questionnaire included 14 questions in four categories: part 1 was about organisation and content of their presentation (questions 1-6); part 2 was related to the use of language (questions 7–9); part 3 was concerned with the manner (questions 10–12); and part 4 (questions 13 and 14) was about the interaction with the audience.

3.4. Weir's analytic speaking criteria

These criteria were modified from Weir (1993). In the current study, these criteria were applied in order to assess the students' speech. These criteria involved five factors which should be considered for assessing the oral performance. These factors were fluency, pronunciation, vocabulary, grammatical accuracy and interactional strategies (Weir, 1993).

3.5. Pre-speaking strategies (Cohen, 1996)

The instrument contains 16 items. The instrument was adopted and adapted from Cohen (1996). The original instrument has three sections: before you speak, while you are speaking and after you speak. The details of the classification were compiled by three teachers. For this study, only *before you speak* was employed. The questionnaire was given in the pre-test and post-test phases. Eight students were selected randomly before giving pre-test to answer the questionnaire for the purpose of piloting the instrument. The reliability of the questionnaire base on Cronbach's alpha was 0.84 (Moradi & Talebi, 2018).

4. Results

Based on the results of QPT, 60 homogenous students were selected and were divided into 4 experimental groups: female peer-assessment group, female self-assessment groups, male self-assessment group and male peer-assessment group (Group 1 = 18, Group 2 = 18, Group 3 = 12 and Group 4 = 12). Groups 1 and 2 became experimental group A and groups 3 and 4 became experimental group B.

Table 1. The results of one-way analysis of variance (ANOVA) among the four different groups

	Sum of squares	df	Mean square	F	Sig.
Between groups					
Within groups					
total			7.139		
287.444					
294.583	3				
56					
59			2.380		
5.133			0.464		0.709

As indicated in Table 1, the sig. level between four experimental groups were 0.709 ($0.709 > 0.05$). Thus, it means that there were not significant differences between four groups based on their OPT scores, and all experimental groups were at the same proficiency level. In this study, we examined whether there is a significant difference between their effects and peer assessment on the performance of English language learners in accurate speech.

Table 2. Independent sample *t*-test for the pre-test scores of self- and peer-assessment experimental groups

		Levene's Test for equality of variances		T-test for equality of means						
		F	Sig.	t	df	Sig.(2-tailed)	Mean difference	Std. Error difference	95% confidence interval of the difference	
								Lower		Upper
<u>Pre- speech scores</u>	Equal	1.96	0.166	176	58	0.861	0.133	0.757	-1.38	1.64
	variances assumed									

As shown in Table 2, the sig. level between the pre-test scores of the two experimental groups (peer and self) was 0.861 ($0.861 > 0.05$). So, it can be concluded that the two experimental groups were at the same level before the treatment. After the treatment and in order to compare the effects of two types of assessment (peer- and self-assessment on the performance in speech accuracy of the students in two experimental groups), an independent sample *t*-test was run.

Table 3. Independent sample *t*-test for the post-test scores of self- and peer-assessment experimental groups

	Levene's Test for equality of variances		t-test for equality of means						
	<i>F</i>	Sig.	<i>t</i>	df	Sig.(2-tailed)	Mean difference	Std. Error difference	95% confidence interval of the difference	
								Lower	Upper
Post-speech scores	0.0	1.00	-3.36	40	0.001	-2.00	0.59	-3.19	0.85
	Equal variances assumed								

As shown in Table 3, the sig. level for the effect of assessment types on post-test scores was 0.001 (0.001 < 0.05). With this in mind, assessment types had a significant effect on the learners' scores. Moreover, students in the experimental peer-assessment group had better performances in comparison to the students in experimental self-assessment group.

In the present study, we examined whether there is a significant difference between male and female English language learners in terms of self-assessment and peers.

Table 4. Descriptive statistics related to post-test scores of participants based on their gender in terms of assessment types

	Assessment types	<i>N</i>	Mean	Std. Deviation	Variance
Post-test	Female	30	15.80	2.64	6.9
	Male	24	16.00	2.37	5.6

As shown in Table 4, the mean scores and SD for female students in the post-test were 15.80 and 2.64 and were 16.00 and 2.37 for male students, respectively. It can be seen that the means were close together, but the mean score for the male students was a little higher than that of the female students. In addition, a two-way ANOVA was run.

Table 5. Two-way ANOVA on male and female students' post-test scores in term of assessment types (peer & self)

Source	Type III sum of squares	df	Mean square	<i>F</i>	Sig.
--------	-------------------------	----	-------------	----------	------

Self and peer	56.01	1	56.01	10.19	0.002
Male and female	0.000	1	0.000	0.000	1.00
Self, peer*male, female	0.278	1	0.278	0.051	0.823

According to Table 5, the sig. level for the effect of assessment types was .002 (sig. = 0.002 < 0.05). Thus, it can be concluded that the differences between the means of the different assessment groups are significant at .05. Sex is not significant because the sig. level is 1.00 (1.00 > 0.05). In addition, there were two independent variables that might mutually affect the dependent variable. This was known as the interaction between the variables and is shown in Table 5 as self, peer*male, female, which is not significant. The calculated significance was 0.823, which is greater than 0.05. Therefore, generally it can be concluded that there were not any significant differences between male and female EFL learners in speech, with regard to self- and peer-assessments.

5. Discussion

The results of this study showed that self-assessment group had lower mean scores in the post-test than that of the peer-assessment group. In other words, the peer-assessment group outperformed the self-assessment group. Post-test data analysis revealed that the type of assessment had a significant effect on the results. According to Yamini and Jafarpoor (2016), self-assessment and peer-assessment are believed to be in line with current learner-centred education. In addition, according to AbolfazliKhonbi and Sadeghi (2012), 'such a finding can be attributed to the introduction between self and peer-assessment, which were new to the students in their associated groups, promoting self-regulatory learning and autonomy (e.g., Dann, 2002; Oscarson, 1989, 1997; Paris & Paris, 2001) and leading to the claim that learning advances from assessing one's own and others' performances' (Stone Wiske, 1999). The present study is in line with the previous research (Ahangari et al., 2013; Ariaifar & Fatemipour, 2013; AbolfazliKhonbi & Sadeghi, 2012; Cheng & Warren, 2005; Colbran, 2017; Khonbi & Sadeghi, 2012; Liang et al., 2020; Lee et al., 2021; Meihami & Varmaghani, 2013; Yamini & Jafarpoor, 2016) and is inconsistent with the Javaherbakhsh (2012) and Ghen's (2014) studies.

Yamini and Jafarpoor (2016) also believed that in comparison with self-assessment, peer-assessment is more influential in terms of its effect on students' course achievement scores. The same results were obtained in a study by Chang et al. (2012), who found that the peer-assessment group outperformed the self-assessment and the teacher-assessment groups. Similarly, Chang et al. (2012) and Sadler and Good (2006) reported that peer-raters are stricter than self-raters. 'Peer-evaluation of writing has also been found to have a significant impact on the improvement of the student writers' (Brown, 2001).

The findings of a study by Sadler and Good (2006) supported the findings of the present study. Their findings indicated that peer-based scores were lower than self-based scores. In other words, it may mean that 'peer-raters tend to under-grade while self-raters tend to over grade'. In better ways, it can be said that peer-raters are stricter than self-raters. Thus, it will be a good reason for better speech performance in peer-group than that of the self-group.

As a final remark, since both peer and self-assessment will assist students 'to map their knowledge and their peer knowledge of the language at various points within a course and/or semester', it seems that taking advantage of these two types of assessments will have better results than those old assessments

like the teacher assessment in language learning and teaching. They also assist students in the development of critical faculties (Coombe & Canning, 2002). These types of assessment may also help students 'look at language in more concrete terms'. Through the use of peer-assessment, students will be motivated to find their classmates' strengths and weaknesses. So, they become more autonomous learners and feel they are real teachers, which can be the main part of the learning process.

References

- AbolfazliKhonbi, Z., & Sadeghi, K. (2012). The effect of assessment Type (self- vs. peer vs. teacher) on Iranian University EFL students' course achievement. *Language Testing in Asia*, 2(4), 47–74. <https://doi.org/10.1186/2229-0443-2-4-47>
- Ahangari, S., Rassekh-Alqol, B., & Akbari, L. (2013). The effect of peer assessment on oral presentation in an EFL context. *International Journal of Applied Linguistics and English Literature*, 2(3), 2200–3452. <https://doi.org/10.7575/aiac.ijalel.v.2n.3p.45>
- Allen, M., Mabry, E., Mattrey, M., Bourhis, J., Titsworth, S., & Burrell, N. (2006). Evaluating the Effectiveness of Distance Learning: A Comparison Using Meta-Analysis. *Journal of Communication*, 54(3), 402–420. <https://doi.org/10.1111/j.1460-2466.2004.tb02636.x>
- Ariafar, M., & Fatemipour, H. R. (2013). The effect of self-assessment on Iranian EFL learners' speaking skill. *International Journal of Applied Linguistics & English Literature*, 2(4), 7–15. <https://doi.org/10.7575/aiac.ijalel.v.2n.4p.7>
- 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*, 37(5), 513–533. <https://doi.org/10.1080/02602938.2010.549204>
- Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy* (2nd ed.). San Francisco, CA: San Francisco State University: Longman.
- Chen, Y. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research*, 12(2), 235–262. <https://doi.org/10.1177%2F1362168807086293>
- Colbran, S. (2017). *Evaluation of the usefulness of self-assessment, peer assessment and academic feedback mechanisms*. Perth, Australia: Curtin University.
- Coombe, C. H., & Canning, C. H. (2002). Using self-assessment in the classroom. *Rationale and Suggested Techniques*, 2(6), 1225–1231. <https://doi.org/10.1002/9781118411360.wbcla046>
- Dann, R. (2002). Assessment as learning: blurring the boundaries of assessment and learning for theory, policy and practice, *Assessment in Education: Principles. Policy & Practice*, 21(2), 149–166, <https://doi.org/10.1080/0969594X.2014.898128>
- Geranpayeh, A. (2003). A quick review of the English Quick Placement Test. *Research Notes*, 12, 8–12. Retrieved from: <https://www.cambridgeenglish.org/Images/23127-research-notes-12.pdf>
- Gunduz, N., Uzunboyulu, H., Ozcan, D. (2017). Developing and testing a scale to assess teachers' attitudes toward peer mediation of student disputes. *Social Behavior and Personality: an international journal*, Volume 45, Number 10, 2017, pp. 1745–1760(16). <https://doi.org/10.2224/sbp.6798>

- Moghimia, A. (2022). On the comparative impact of self-assessment and peer assessment on Iranian male and female EFL learners' accuracy in speech. *Contemporary Educational Researches Journal*, 12(4), 204–213. <https://doi.org/10.18844/cerj.v12i4.7468>
- Hooshangi, R., Yousofi, N., & Mahmoudi, O. (2014). Self-Assessment on Iranian foreign language learners' oral performance ability: A case study. *Iranian-Rehabilitation-Journal*, 12(4), 65–71. Retrieved from: <http://irj.uswr.ac.ir/article-1-350-en.html>
- Housen, A., & Kuikken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Appl Linguist*, 30(4), 461–473. <https://doi.org/10.1093/applin/amp048>
- Javaherbakhsh, M. R. (2012). The impact of self-assessment on Iranian EFL learners' writing skill. *English Language Teaching*, 3(2), 111–112. <https://doi.org/10.5539/elt.v3n2p213>
- Kargar, A. A., Salimi, M., & Zareian, A. (2014). Self assessment of pronunciation and its implication on EFL learners' pronunciation ability. *International Journal of Language Learning and Applied Linguistics World*, 36(4), 506-516. Retrieved from: <https://eric.ed.gov/?id=EJ819325>
- Lee, H., Fridlind, A. M., & Ackerman, A. S. (2021). An evaluation of size-resolved cloud microphysics scheme numerics for use with radar observations. Part II: Condensation and evaporation. *J. Atmos. Sci.*, 78(5), 1629-1645. <https://doi.org/10.1175/JAS-D-20-0213.1>
- Li, J., Huang, J., & Cheng, S. (2022). The reliability, effectiveness, and benefits of peer assessment in college EFL speaking classrooms: Student and teacher perspectives. *Studies in Educational Evaluation*, 72, 101120. <https://doi.org/10.1016/j.stueduc.2021.101120>
- Liang, H. Y., Tang, F. I., Wang, T. F., & Yu, S. (2021). Evaluation of nurse practitioners' professional competence and comparison of assessments using multiple methods: Self-assessment, peer assessment, and supervisor assessment. *Asian Nursing Research*, 15(1), 30–36. <https://doi.org/10.1016/j.anr.2020.10.004>
- Lindsay, C., & Knight, P. (2006). *Learning and teaching English: A course for teachers*. London, UK: Oxford University Press Oxford.
- Luoma, S., & Tarnanen, M. (2003). Creating a self-rating instrument for second language writing: From idea to implementation. *Language Testing*, 20(4), 440-465. <https://doi.org/10.1191/0265532203lt267oa>
- Meihami, H., & Varmaghani, Z. (2013). The Implementation of self-assessment in EFL writing classroom: An experimental study. *International Letters of Social and Humanistic Sciences*, 9(2), 39–48. doi:[10.18052/www.scipress.com/ILSHS.9.39](https://doi.org/10.18052/www.scipress.com/ILSHS.9.39)
- Moradi, F., & Talebi, A. (2018). Providing a management model to evaluate the utility area for the use urban runoff in the management of urban green spaces (Case study: KARAJ City). *Journal of Environmental Science and Technology*. <https://doi.org/10.22034/jest.2018.10916.1946>
- Mousavi, S. A. (2012). *An encyclopaedic dictionary of language testing*. Tehran, Iran: Rahnama Press.
- Oscarson, M. (1989). Self-assessment of language proficiency: rationale and applications. *Research Article*, 6(1), 1-13. <https://doi.org/10.1177/026553228900600103>
- Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36(2), 89–101. https://doi.org/10.1207/S15326985EP3602_4
- Patri, M. (2002). The Influence of Peer Feedback on Self- and Peer-Assessment of Oral Skills. *Language Testing*, 19, 109-131. <https://doi.org/10.1191/0265532202lt224oa>
- Richards, J. C., & Schmidt, R. (2002). *Dictionary of language teaching and applied linguistics* (3rd ed.). London, UK: Pearson Education Limited.

- Moghimia, A. (2022). On the comparative impact of self-assessment and peer assessment on Iranian male and female EFL learners' accuracy in speech. *Contemporary Educational Researches Journal*. 12(4), 204–213. <https://doi.org/10.18844/cerj.v12i4.7468>
- Sadler, P. M., & Good, E. (2006). The Impact of Self and Peer-Grading on Student Learning. *Educational Assessment* 11(1), 1-31. https://doi.org/10.1207/s15326977ea1101_1
- Saito, H. (2008) EFL Classroom Peer Assessment: Training Effects on Rating and Commenting. *Language Testing*, 25, 553-581. <http://dx.doi.org/10.1177/0265532208094276>
- Sluijsmans, D. (2002). *Student involvement in assessment. The training of peer assessment skills* (Unpublished doctoral dissertation). Heerlen, Netherlands: Open University of the Netherlands.
- Van Loon, J. R. (2007). The Oxford Quick Placement Test, pilot report. Retrieved from: <https://studylib.net/doc/9049250/the-oxford-quick-placement-test-questionnaire>
- Weir, C. (1993). *Understanding and developing language tests*. New York: Prentice Hall.