Involving Self-, Peer-, and Teacher-Assessment in a Writing Course: A Cross Comparison and Reflection of Students' Perceptions

Gholam-Reza Abbasian

Assistant Professor of TEFL, Department of English Language and Literature, Imam Ali University, Tehran, Iran.
E-mail: gabbasian@gmail.com

Maryam Khezrinejad*(Corresponding Author)
Lecturer in TEFL, Department of English,
Islamic Azad University, Islamshahr Branch, Islamshahr, Iran.
E-mail: khezri53@gmail.com

Mojtaba Teimourtash

Ph.D. Candidate of TEFL, Department of English,
Islamic Azad University, Science and Research Branch, Tehran, Iran.
E-mail: mojtabateimourtash@gmail.com

Abstract

The increasing demand for lifelong learners and reflective practitioners has re-conceptualized the connection between assessment and learning to the extent that alternative assessment methods (i.e., self-, peer and teacher-assessment, etc.) have emerged. However, their incorporation into various language skills might bring about certain consequences. Among them, the writing skill is often perceived as unique in its nature in terms of both teaching practices and assessment modalities. In a bid to exercise a Learner-Oriented Assessment (LOA) practice, the present study was designed to comparatively implement self-, peer-, and teacher-assessments in a writing course, and explore the experiences and perceptions of the learners towards the three assessment alternatives. Pertinent statistical analyses revealed significant differences among them such that higher proficiency level learners evaluated their writings more realistically, while lower proficiency level overestimated their abilities. Moreover, most of them had positive attitudes towards this novel assessment experience, holding that LOA could scaffold them in gaining a lot and arousing their awareness of their weaknesses and strengths.

Introduction

Assessment is regarded as a critical part of any instructional process. It is increasingly gaining acceptance that it is significant for both learners and teachers to get involved in assessment and to have control over its methods, processes, results, and the underlying rationale (Cheng & Waren, 2005). Jafarpur (1991) also notifies that to increase learners' responsibility in foreign language study programs, the modification of testing procedures would be necessary.

As a result of the growing focus on learner independence and autonomy, LOA has gained much attention. As LeBlanc and Painchaud (1985) emphasize, the learners need to have an active involvement in the process of learning; and as assessment is a basic component in the educational process, this active participation also includes contribution in assessment.

In higher education, especially, the assessment of student learning has experienced a move from traditional testing of knowledge towards assessment of learning (Dochy, Segers, & Sluijsmans, 1999). The paradigm shift in language testing practice from psychometrics to educational assessment, and the increasing criticism over standardized tests, especially in light of the educational reform movements, encouraged the educators to use some more indirect, formative, and holistic approaches to assessment, instead of applying the standardized summative techniques.

Therefore, there was a need to apply new assessment approaches which were more multiculturally sensitive and were free from linguistic, normative, and cultural biases dominating the traditional testing systems in order to come up with equity in education and to achieve educational excellence for all groups of students (Hamayan, 1995; Huerta-Macias, 1995; Martin-Kniep, 2000; Soodak, 2000).

Unlike testing culture, which aims at measuring factual knowledge and low-level cognitive skills, an assessment culture seeks to assess higher-order thinking processes and competencies (Birenbaum & Dochy, 1996). The aim of assessment culture is to align assessment with instruction and to provide learner with sufficient opportunities to receive feedback from their learning. Students should also have an active role in the learning and assessment processes. To this end, students would need to possess required skills to regulate their studying and to reflect on their learning outcomes and practices (Lindblom-Ylänne, Pihlajamäki, & Kotkas, 2006). By the same token, students would also need to develop necessary strategic learning behaviors to choose the most effective learning strategies and practices so that they can effectively deal with the demands of their learning environments (Biggs, 1999; Lindblom-Ylänne & Lonka, 2001). In this regard, Ahmadpour and Yousefi (2016) focus on the advantages of group-focused learning and assert that teachers would need to help learners acquire the learning strategies that are required in coping with future activities. In the recent up-to-date language testing articles, this new trend in assessment is known as the 'alternative assessment movement' (Worthen, 1993; Bachman, 2000; Alderson & Banerjee, 2001).

Alternative assessment

Like many other terms and concepts in the field of language teaching and testing, there is no single undisputed definition of 'alternative assessment' in literature. Some educators look at alternative assessment as a reaction to standardized methods of assessment such as the objective tests consisting mostly of professionally-prepared multiple-choice items (Huerta-Macias, 1995). Some others view alternative assessment as more general terms. For instance, Hamayan (1995) defines alternative assessment as "procedures and techniques which can be used within the context of instruction and can be easily incorporated into the daily activities of the school or classroom" (p. 213). In the same vein, Smith (1999) maintains that "alternative assessment might take place outside the classroom or even the institution at various points in time, and the subjects being tested may be asked to present their knowledge in various ways" (p. 703).

Kohonen (1997) uses the term 'authentic assessment' instead of alternative assessment and defines it as follows:

> … the procedure for evaluating learner performance using activities and tasks that represent classroom goals, curricula and instruction in as realistic conditions of language use as possible. It uses such forms of assessment that reflect student learning. It emphasizes the communicative meaningfulness of evaluation and the commitment to measure that which we value in education. It uses such forms of assessment that reflect student learning, achievement, motivation and attitudes on instructionally-relevant classroom activities ... Its results can be used to improve instruction, based on the knowledge of learner progress. (p. 6)

In the same vein, Alderson and Banerjee (2001) later provide the following definition:

> Alternative assessment is usually taken to mean assessment procedures which are less formal than traditional testing, which are gathered over a period of time rather than being taken at one point in time, which are usually formative rather than summative in function, are often low-stakes in terms of consequences, and are claimed to have beneficial washback effects. (p. 228)

Two types of alternative assessments which are also practiced in the present research are regarded as self-assessment and peer-assessment; in the following sections, each will be discussed in details.

Self-assessment

Self-assessment, which is also referred to as self-rating, self-report, self-appraisal, self-control, self-evaluation, self-estimate, or self-testing is a good example of formative assessment/learning which is seen as "an internal or self-directed activity"

(Oscarson, 1989, p. 1); it is a process where students evaluate their own product or performance against a standard.

Self-assessment, rather than replacing teacher assessment, mostly supplements it. The main characteristic of self-assessment is that it involves students in specifying standards and criteria to apply to their work, and determining the extent to which they have met these criteria and standards. Although self-assessment is commonly a supplement to teacher assessment of students, in some cases it may even replace it (Habeshaw, Gibbs, & Habeshaw, 1995).

According to Spiller (2012), involving students in establishing criteria for self-assessment tasks can help them develop their understanding of what constitutes quality products in a specified area. Moral reasons such as power sharing between teacher and learner as well as motivational reasons, including individuals' excitement of self-discovery, are attractive justifications for self-assessment practices against the accusations which are raised regarding the lack of reliability of such practices.

Focusing on the formative aspect of self-assessment, Andrade and Du (2007) refer to self-assessment as "a process of formative assessment during which students reflect on and evaluate the quality of their work and their learning, judge the degree to which they reflect explicitly stated goals or criteria, identify strengths and weaknesses in their work, and revise accordingly" (p. 160).

Self-assessment can make students aware of their learning progress. It can also motivate further learning, encourage reflection on learning, develop necessary skills for lifelong learning, and promote learner responsibility and independence (Spiller, 2012). The use of self-assessment can encourage the direct involvement of learners in learning process and can promote the integration of cognitive abilities with affective learning (Hart, 1994; O'Malley & Valdez Pierce, 1996; Kohonen, 1997).

For self-assessment to be meaningful, as Kohonen (1997) maintains, it is important to include an essence of learner choice in the learning process; to do this, at least a negotiated, partially open curriculum needs to be employed (e.g. through learner-initiated and monitored project work). Kohonen adds that, in this process, the learning contracts are negotiated in class, and the teacher explains the evaluation criteria to learners so that they can have a proper idea of acceptable learning outcomes.

Different methods of self-assessment

Self-assessments, according to Brown and Hudson (2002), require students to rate their own language through one of the following methods:

- *Performance self-assessments* require students to read a situation and to evaluate how well they would respond in that situation.

- *Comprehension self-assessments* require students to read a situation and decide how well they would comprehend that situation.

- *Observation self-assessments* require students to listen to audio- or watch video-tape recordings of their own language activities/performance which are usually recorded in natural situations such as in role-play activities, and then to decide how well they think they have performed.

The Rationale of self-assessment

Self-assessments are a way of attaining information about one's proficiency in a language. As LeBlanc and Painchaud (1985) have pointed out, "it has now become a commonplace belief that to be efficient, a teaching/learning strategy requires that students have some input in the complete learning cycle" (p. 673). This means that the role of students in the process of teaching and learning should not be limited to the study of the subject matter determined by others and the use of techniques enforced by others, but they need to get involved in the process. Here, there are two points which are worth noting about self-assessment.

Self-assessment is not the same as self-grading. As Andrade and Du (2007) postulate, during the process of self-assessment, learners review their work to determine the extent to which it reflects the goals of the assignment as well as the assessment criteria later used by the teacher in evaluating their work. They generally view "self-assessment as feedback for oneself from oneself" (p. 160), and declare that utilizing any method of self-assessment can assist students in developing a kind of ability to evaluate their own practice objectively and then modify their work and enhance its quality.

Mousavi (2011) briefly presents the rationale of self-assessment as follows:

- *To promote learning:* It helps students learn how to evaluate, which is quite advantageous in learning. Students' ability to make reliable and trustworthy evaluations/judgments on the effectiveness of their practice is a crucial part of learning process.

- *To raise awareness level:* Getting involved in the process of self-assessment, students will be encouraged to consider course content and assessment principles in a more cognizant way and to improve their evaluative attitudes on the 'whats', 'hows', and 'whens' of assessment.

- *To improve goal orientation*: Self-assessment can improve learners' knowledge of potential goals in most language learning contexts while other-directed assessment often makes the learner accept the test goals and testing techniques.

- *To expand the range of assessment:* As in many respects, learners usually over-estimate their language competence; learner involvement in the assessment process can provide them with a broader view of real language ability and communicative competence.

- *To share the burden of assessment:* Involving learners in assessment process can decrease the assessment burden which is usually on the teacher

and can make the teacher somewhat free for other responsibilities such as final grading and the like which cannot be undertaken by the learners.

- *To enhance post-course effects:* Self-assessment practice can enhance autonomous learning which is regarded as a significant objective in language learning (p. 663).

Advantages of self-assessment

Andrade & Valtcheva (2009) view self-assessment as a valuable learning instrument and a part of an assessment process. They believe that through self-assessment, students can take the following advantages:

- They can recognize their own skill gaps, as well as the weaknesses in their knowledge

- They can specify those parts of learning process where they need to focus their attention

- They can set more realistic goals

- They can modify their practice

- They can trace their progress

- If online, they can decide when to move to the next level of the course

Mousavi (2011) also considers some advantages such as direct involvement of students, greater learner autonomy, and increased motivation to learn.

Limitations of self-assessment

Despite several advantages and benefits, several shortcomings and disadvantages have been conceived of for self-assessment. A common misconception about self-assessment, as Andrade and Du (2007) claim, is that students' grading themselves cannot be taken much seriously. Some consider assessment as among teacher's responsibilities which should not be carried out by the students. Unfortunately, there are some research findings which reinforce such arguments. For instance, Eva & Regher (2005) reviewed several studies questioning the usefulness of self-assessment, as there were doubts over the students' ability to rate their own strengths and weaknesses.

There are also concerns regarding the accuracy (i.e. agreement between self and teacher marks) and consistency (i.e. scores being produced repeatedly) of students' self-assessment results (as in Eva & Regehr, 2005; Ross, 2006; Lew, Alwis, & Schmidt, 2010), so that Lew et al. (2010) consider the accuracy of student self-assessment as weak to moderate. Although there are some other studies which have clearly come up with the benefits (e.g. Rolheiser & Ross, 2000; Cyboran, 2006), there are still misconceptions regarding the usefulness of self-assessment.

Blatchford (1997) found that the success of self-assessment depended on the subject matter and the age of student age. Ross (2006) reported inconsistencies

between self-assessments when they were conducted at different times. Therefore, the reliability of self-assessment practice in terms of accuracy and consistency remains doubtful.

In a few words, the advantage of self-assessment may not be judged in terms of accuracy and consistency of assessment; literature on self-assessment also agrees that the accuracy and consistency of assessment are not the critical attributes to students' learning (Eva & Regehr, 2005). Instead, the important point is that self-assessment process encourages the students to critically evaluate their own work seeking to improve themselves (Andrade & Du, 2007), and this is the influential factor that can contribute to learning (Willey & Gardner, 2010).

Peer-assessment

Peer assessment includes the assessment process in which students assess the achievements, learning outcomes, and performances of their peers. According to Topping (2009), "peer assessment is an arrangement for learners to consider and specify the level, value, or quality of a product or performance of other equal-status learners" (p. 20). Topping believes that peer-assessment can be summative or formative, and a variety of products including writing, oral presentations, portfolios, test performance, or other skilled behaviors can be assessed by peers.

In the process of peer-assessment, students will learn how to make self-determining judgments while commenting on and judging the work of other students. Through applying a variety of assessing methods and undertaking group activities, peer-assessment can also develop teamwork skills in students and can enable them to undertake a wider variety of tasks than could be done by a single individual. Though peer assessment plays a significant role in formative assessment, if carefully implemented, it can also be a part of summative assessment (Habeshaw, et al., 1995).

Boud and Falchikov (2007) maintain that "peer assessment requires students to provide either feedback or grades (or both) to their peers on a product or a performance based on the criteria of excellence for that product or event which students may have been involved in determining". Topping (1998) also defines peer assessment 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" (p. 250). In the process of peer assessment, students grade and/or give feedback on the work of their peers, and are also judged for the quality of the evaluations they have made (Davies, 2006).

Patri (2002) believes that if learners can assess the quality of their own achievement and level of their own performance as well as those of their peers, it is prospective that they will be able to grasp and follow the assessment criteria. According to Brown (2004), both self and peer assessment involve students in their own destiny, promote learner autonomy, and stimulate motivation. Cheng and Warren (2005) also view it as being crucial for both teachers and students to involve in and control over the methods, procedures, and outcomes of assessment, and to understand the underlying rationale for assessment. The main pedagogical value of

peer assessment is that it allows learners to play a part in the assessment process and provides learners with opportunities to have an active role in the evaluation of their peers' learning process and products (Peng, 2010).

Advantages of peer-assessment

Many scholars have counted some important benefits of peer-assessment. For instance, Nilson (2003) believes that it can improve critical thinking, lifelong learning, communication, and collaborative skills. Topping (1998) notes that it improves the amount of feedback and enhances higher order thinking. Sivan (2000) views it as an instrument which can enhance sense of ownership and responsibility in learners, and can promote active autonomous learning. Spiller (2012) believes that it can reduce power imbalance between students and teachers and can heighten the status of students in the leaning process. In addition, a study conducted by Williams (1992) indicate that most students view both self- and peer-assessment useful, interesting, and fun.

Chen and Warren (2005) hypothesize that, in the process of peer-assessment, learners can develop several types of skills and abilities that cannot be provided to them when their works are assessed by the teacher alone. They also assume that peer-assessment can also provide learners with the chance to analyze, monitor, and evaluate their peers in term of both process and product of their learning.

Research indicates that peer-assessment also raises higher order cognitive thought and higher order reasoning in students (Birdsong & Sharplin, 1986), assists student-centered learning among undergraduate learners (Oldfield & MacAlpine, 1995), boosts active and flexible methods of learning (Entwhistle, 1993), allows students to have a deeper approach to learning (Entwhistle, 1987; 1993; Gibbs, 1992), and, acting as a socializing force, develops interpersonal skills and relationships between students as it (Earl, 1986).

Topping (1998) believes that peer-assessment of writing has positive effects on the achievement and attitudes of students. Moreover, Cho and MacArthur (2010) report that receiving feedback from peers can provide learners with more opportunities to develop their writing quality compared to when they receive feedback from a teacher only. They also found that cognitive feedback such as clarification and explanation of problems was more useful for students' learning than affective feedback such as praising or criticizing.

Another study by Cho and Cho (2011) also indicates that giving feedback to peers can cause more improvements in writing skills than receiving feedback. Boud and Falchikov (2007, p. 139) also see peer involvement in assessment where particular characteristics are present. These include features which:

- are planned to enhance learning;

- need learners to take responsibility for their behaviors;

- stimulate a reflective approach to learning;

- need learners to specify and use standards and criteria;

- create some sort of modelling and/or scaffolding;

- let learners practice peer- and self-assessment skills in a diversity of contexts;

- let the removing of support so that learners can get closer to the assessment autonomy.

Limitations of peer-assessment

One of the problems with peer-assessment is that the students are sometimes reluctant to criticize their friends and receive subjective marks. Cheng and Warren (2005) report that when assessing peers, students are usually uncomfortable and insecure due to their own perceptions of inability. The research carried out by Orsmond, Merry, and Reiling (1996) also shows the same result. Another shortcoming assumed for peer assessment is that due to the need for training, preparing, and monitoring the students, it is quite time-consuming (Topping, 1998; Cheng & Warren, 2005). Peer-assessment is also criticized for not being objective. To achieve objectivity, Freeman (1995) asserts that the students should be appropriately trained, and should extensively practice in peer assessment. Patri (2002) also agrees with this idea and puts emphasis on the role of training and experience in enhancing the objectivity of peer-assessment practices. Brown (2004) also agrees that subjectivity is the major weakness of peer-assessment and needs to be removed. Sometimes, students may either be too critical or too boastful on themselves. In addition, they might not know how to make adequate judgments; they might also feel anxious and resistant toward peer-assessment, especially in the beginning stages (Topping, 1998).

Peer-assessment in EFL contexts

Reviewing the related literature in EFL context indicates that peer-assessment has been more typically integrated into English language writing where peers provide feedback to and/or edit their peers' written work with the aim of helping with revision (for example, Birdsong & Sharplin, 1986; Devenney, 1989; Bell, 1991; Mangelsdorf, 1992; Mendonca & Johnson, 1994).

According to Langan and Wheater (2003), the way peer-assessment schemes are planned and managed has a crucial role in determining their success. Several scholars such as Race (1999), Magin and Helmore (2001), and Stefani (1994) have suggested related guidelines for the management of peer-assessment. In general, they suggest that peer-assessment systems should keep everybody involved in the process (e.g. about why and how to allocate the marks); should have a simple, highly objective assessment system; should, whenever possible, negotiation advance the assessment criteria with students; should have the teachers second-mark some of the assessments done by the students; should provide students with the opportunity to discuss and/or challenge the allocated marks; and should provide some forms of

feedback to students in order to approve that the marks assigned by peers are valid and not much different from those assigned by the teachers.

Self-, peer-, and teacher-assessment

Self- and peer-assessment are not limited to students grading their own or a peer's productions, but they also let them take a role in determining what high-quality learning means in a specific case (Brown, Bull, & Pendlebury, 1997; Dochy et al., 1999; Topping, 2003). Both self- and peer-assessment, according to Lindblom-Ylänneet and Lonka (2001), can be considered as learning tools for the development of different skills such as those required for professional responsibility, judgment, and autonomy.

Topping (2003) holds that self- and peer-assessment can be either summative or formative: they are summative when the goal is to judge learning results or to assign quantitative grades to students, while they are regarded as formative when they focus on the qualitative assessment of different kinds of learning results (Topping, 2003). However, it is believed that peer-assessment should be formative in nature because summative peer-assessment can decline cooperation between students (Boud, 1995).

However, regarding the accuracy and reliability of scores, teachers' ratings are usually considered as the reference point, compared to self- and peer-assessment (Topping, 2003). However, the reason for such difference in scoring is that there are different understandings of assessment criteria by teacher and students (Orsmond et al., 1996, 1997, 2000, 2002). It seems that more proficient students have a tendency to underrate their performance, whereas less proficient ones tend to overvalue their performance (Dochy et al., 1999; Lejk & Wywill, 2001).

In this regard, Dochy et al. indicate that self-assessment skills develop as the learners improve their proficiency level, because advanced students seem to predict their performance better than novices. Moreover, critical analysis of one's performance appears to be more difficult than evaluating a peer's performance in a group (Segers & Dochy, 2001). However, research shows that compared with teacher-assessments, students usually overestimate both their capabilities and their performance (Zoller & Ben-Chaim, 1997).

Although there are usually some briefing sessions before starting the assessment process, it seems that students and teachers still have different understandings of individual assessment criteria (Orsmond et al., 1996, 1997, 2000); this suggests that further research is necessary to explore these types of assessment practices (i.e. self-, peer-, teacher-assessment) as deeply as possible.

Purpose and problem

Concerning the need to delve more into different types of alternative assessment, the present research tries to elucidate and compare self-, peer-, and teacher-assessment in the marking practice of language learners' essays. The present study aimed that implementing LOA in a writing course in order not only to compare the three

scoring modes (i.e. self-, peer-, and teacher-assessments) but to scrutinize the experiences and perceptions of the learners towards such experience. In other words, the researchers made an effort to discover whether there is any difference among the three scoring modes and to investigate the participants' attitude towards this relatively different experience.

Apart from the capabilities of students in using the rubrics and scoring the writing practices, it seems undeniable to make completely sure that the students put the same criteria into practice without considering idiosyncratic differences towards implementing such rubrics. However, it is crucial to note that fading out the role of students' attitudes towards such novel experiences can alleviate the robustness of the desired outcomes; therefore, taking the role of students' attitudes into their scoring practices seems to play a pivotal role in crystallizing the significance of both plausibility and feasibility of using different modes of LOA.

Method

Participants

The participants of the study were 104 male and female sophomore university students majoring in English translation. The participants were at different levels of proficiency ranging from lower intermediate to upper intermediate because they had entered the university without passing any entrance exam, or any sort of placement exams, to make them more homogenized; therefore, they had weaknesses and strong points in different aspects of language. There were also two experienced university instructors who rated the writings and coached the students to work based on the rubric.

Materials

The research was carried out in three essay-writing classes whose students had already passed a paragraph writing course. Every week, each student needed to write an essay and deliver it to the teacher for further comments and feedbacks. The type and subject of writings were changing each week, so there was no single type of writing worked on during the research.

Procedures

According to O'Malley and Pierce (1996), "self-assessment is a process through which students must be led" (P. 39); it is not about forms or checklists only, but about teaching, or in better terms coaching, students to realistically evaluate their progress and to learn new skills including proper assessment practices, accurate writing, and reasonable judgments. In this research, the students were provided with both oral and written instructions about not only how to write but how to grade essays. Apart from teaching academic methods of essay-writing and the associated types, the teachers also spent three sessions on making the students acquainted with the new methods and procedures of assessment—especially LOA which was the primary focus of this study. The origins and background of testing and assessment as well as new advances in such fields were fully explained to the participants so that

they can have a more comprehensive overview of both traditional and modern practices in the field of assessment.

Before any assessment procedure, the teachers planned for and carried out some concise elaborated conversations with the aim of making the students prepared for getting involved in both self- and peer-assessment. Underlying assumptions and principles of LOA innovations were introduced to the students, and the rubrics based on which they needed to carry out the assessment practice was fully elaborated in details. They were also provided with the opportunities to put the assumptions into practice authentically.

As the way of carrying out the self-assessment process is quite crucial (Boud, 1995), the implementation process involved some preplanned phases. The first phase included the presentation of origin, background, rationale, procedures, and outcomes of the LOA, with an overt focus on self- and peer-assessment. Students were involved both in introducing the judgment criteria and in evaluating their own and their peers' works. With regard to the development of assessment abilities, they were coached and assisted by the teachers in an ongoing supportive fashion.

In the second phase, after necessary explanations and practices, the essays assigned as weekly assignments were judged for three consecutive weeks by four people. First, every week each student graded his/her own essay (self-assessment). Second, his/her essays were graded by a peer (peer-assessment). Finally, all essays were graded by two teachers (teacher-assessment) for the purpose of getting assured of inter-rater reliability of scores. Throughout the process, the students were provided with plenty of chances to raise questions and discuss in connection with the criteria and the way it can be used and implemented. The aforementioned procedure is schematically crystallized in the following flowchart:
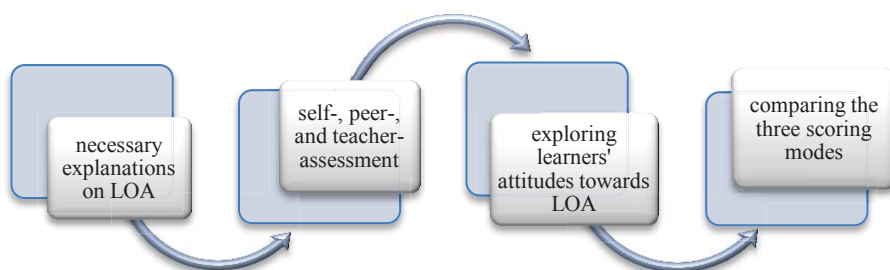
The assessment criteria selected for the purpose of this research was the one suggested by Brown (2007) which is an adapted version of the criteria by J. D. Brown (1991). The rubric was tangible enough for the students so that they could use it conveniently.

Categories for evaluating writing (adopted from J. D. Brown, 1991, pp. 42-46).

Content (scoring 0-24)

- thesis statement
- related ideas
- development of ideas through personal experience, illustration, facts, and opinions
- use of descriptions, cause/effect, and comparison/contrast
- consistent focus

Organization (scoring 0-20)

- effectiveness of introduction
- logical sequences of ideas
- conclusion
- appropriate length

Discourse (scoring 0-20)

- topic sentences
- paragraph unity
- transitions
- discourse markers
- cohesion
- rhetorical conventions
- reference
- fluency
- economy
- variation

Syntax (scoring 0-12)

Vocabulary (scoring 0-12)

Mechanics (scoring 0-12)

- spelling
- punctuation
- citation of references (if applicable)
- neatness and appearance

The teacher-assessments were carried out independently so that none of the raters was aware of the other raters' grades. Teachers were not aware of the results of self- and peer-assessments before grading the essays, and the students did not

assess their peers reciprocally. The scoring matrix included six sections of content, organization, discourse, syntax, vocabulary, and mechanics, each of which tapping one particular facet of writing indeed. The first category could be scored from 0 to 24, the second from 0 to 20 and the last three categories from 0 to 12 (the total score was 100). Finally, the results obtained from, self-, peer-, and teacher-assessments were compared with each other.

After accomplishing the assessment practices, all participants were given a list of questions which they needed to elaborately answer based on what they felt and experienced in the processes of self- and peer-assessment. The aim of raising such questions was to scrutinize the participants' attitudes towards the novel types of assessment practice they had experienced in a qualitative manner. In this research, the students were provided with the opportunity to get involved in the rating process and to get familiar with the related challenges and the skills required.

Data Analysis

As mentioned in the previous section, the scores given by the two teachers were correlated in order to estimate the inter-rater reliability of teachers' scores. As the results showed highly positive correlation, the researchers got confident about the inter-rater reliability of scores; then the mean of scores given by teachers to each student was regarded as the result of teacher-assessment. In the next phase of data analysis, the differences among self-, peer-, and teacher-assessment were compared. The findings are illustrated on the following graph:
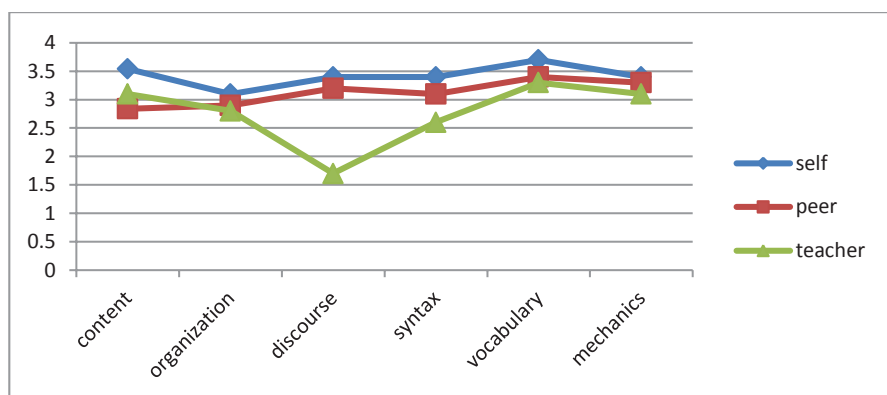

Figure 1. Comparison of self-, peer-, and teacher-assessment

In order to compare the three sets of scores in a more precise manner, ANOVA analyses were also conducted in three stages: between self- and teacher-assessment, between peer- and teacher-assessment, and among self-, peer, and teacher-assessment scores. The results are provided in the following tables:

Table 1. Descriptives

VAR00002

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
self	104	78.8400	10.80506	1.52807	75.7692	81.9108	49.00	94.00
peer	104	75.9400	11.47386	1.62265	72.6792	79.2008	45.00	91.00
teacher	104	71.7400	14.48125	2.04796	67.6245	75.8555	39.00	89.00
Total	312	75.5067	12.61763	1.03023	73.4709	77.5424	39.00	94.00

Table 2. Test of homogeneity of variance

VAR00002

Levene Statistic	**df1**	**df2**	**Sig.**
4.125	2	147	.018

Table 3. ANOVA of the results of self-, peer-, and teacher-assessment

VAR00002

	Sum of Squares	**df**	**Mean Square**	**F**	**Sig.**
Between Groups	1274.333	2	637.167	4.173	.017
Within Groups	22447.160	147	152.702		
Total	23721.493	149			

Table 4. Post-hoc tests of multiple comparisons

Dependent Variable:VAR00002

	(I) VAR00001	**(J) VAR00001**	**Mean Difference (I-J)**	**Std. Error**	**Sig.**	**95% Confidence Interval**	
						Lower Bound	Upper Bound
Bonferroni	self	peer	2.90000	2.47145	.728	-3.0851	8.8851
		teacher	7.10000*	2.47145	.014	1.1149	13.0851
	peer	self	-2.90000	2.47145	.728	-8.8851	3.0851
		teacher	4.20000	2.47145	.274	-1.7851	10.1851
	teacher	self	-7.10000*	2.47145	.014	-13.0851	-1.1149
		peer	-4.20000	2.47145	.274	-10.1851	1.7851
Tamhane	self	peer	2.90000	2.22890	.481	-2.5148	8.3148
		teacher	7.10000*	2.55521	.020	.8840	13.3160
	peer	self	-2.90000	2.22890	.481	-8.3148	2.5148
		teacher	4.20000	2.61288	.298	-2.1530	10.5530
	teacher	self	-7.10000*	2.55521	.020	-13.3160	-.8840
		peer	-4.20000	2.61288	.298	-10.5530	2.1530

*. The mean difference is significant at the 0.05 level.

Results and discussion

Based on the findings illustrated in the above graph, the following conclusions were drawn and might be generalized, to some extent, to other similar circumstances,

especially in Iranian EFL contexts. The conclusions discussed here can be categorized into different groups which, definitely, share overlaps with one another.

First, by analyzing the graph, it is vivid that students gave higher scores in self-assessment phase in comparison with the peer-assessment. It triggers one of the merits of peer-assessment through which learners get involved in the process of learning meticulously when it comes for them to play a role as a scorer. Meanwhile, teachers, since following the clear-cut criteria and considering the table of specification of the course in assessing the performance of the students, scored the students less in almost all areas; however, it seems to be true that it might be more near to the real performance of the students. In case of discourse, for instance, the dramatic difference between scores in self- and peer-assessment phase and ones in teacher-assessment phase is undeniable.

Another point which should be taken into account is that the peer assessment was done, to some extent, haphazardly, i.e. subjectively rather than following an objective fashion, regardless of the training sessions; nonetheless, both self- and peer-assessment practices seems to be quite helpful for the students not only in getting familiar with the scoring criteria but also in learning how to have a critical wider look at both assessment and writing processes as it gets them involved more directly.

It is worth noting that by considering the raw scores meticulously, this point was spotlighted by the researchers that overachievers, those who gained a lot in the writing course, estimated their knowledge of syntax and words more realistically; in simple terms, they did the self- and peer-assessment accurately not to miss any mentioned points or guidelines by their instructors in the classroom settings. On the contrary, underachievers, who did not learn enough during the course, had no crystal clear understanding of their real performance. The results seem to be considered as an overestimation of their actual performance.

Yet, the most important issue is that the Iranian attitudes and approach towards writing, as well as scoring or assessing the performance of the students in writing skill, are not straight enough as what is needed to be in English writing skill. Along the same vein, Boroditsky (2001) goes on to hold that "language is most powerful in determining thought for domains that are more abstract, that is, ones that are not so reliant on sensory experience" (p. 19). More specifically, scholars such as Lay (1982) contend that individuals who are more developed in L1 literacy skills show better performance in second language writing. However, Connor (1996) regards language and writing as cultural phenomenon, and argues that there are some rhetorical conventions which are unique to each language and make individuals' L1 writing distinct from their L2 writing. To put flesh on the issue, Kaplan (1972) maintains that texts written in English are typically of *linear* and *hierarchical* structure, because English speakers are normally inclined to be direct and straightforward in both writing and speech. Due to their prevailing thought patterns, English speakers mostly utilize direct expressions and patterns, while oriental people prefer to use indirect patterns. Kaplan's cultural thought patterns are illustrated in the following patterns:
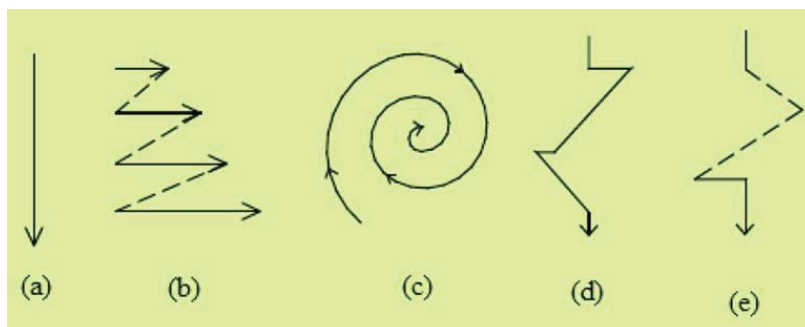
Figure 2. (a) English, (b) Semitic (Hebrew and Arabic), (c) Oriental, (d) Roman, (e) Russian Discourse Structures by Kaplan (1972)

Regarding what was mentioned above, writing must be straight forward in English, but Iranian learners, technically speaking, are culturally different in terms of codes of writing. In other words, Iranian learners write in a more complicated fashion in a way that it seems they are not straightforward at all, i.e. they replicate their thought patterns of writing in Persian into English classes. In the same vein, when the areas of cohesion, coherence, discourse, etc. are concerned, they get violated in the process of writing. Put simply, in both self- and peer-assessments, participants usually follow the same stream of thought of their L1; consequently, they score the culture-bound areas, such as discourse, less than what they are supposed to, even if the rules of English writing style in that very area got violated.

The last but not the least, after deeply reviewing the students' responses to the questions regarding their experiences in judging their own and their peers' essays, the researchers found only positive comments of the students; almost all of the students expressed positive attitudes towards their novel experience and the evaluation process. They found it both motivating and constructive to look at the writing practice from a new critical perspective. Most of the participants expressed that they learnt a lot from the assessment practice and became much more cognizant of their strengths and weaknesses. Many of them regarded both self- and peer-assessment as a tool to learn more about different aspects of essay writing including grammar, content, organization, etc. because for the purpose of judging an essay, having enough knowledge about organization of a good piece of writing, the elements involved and the points to be considered are a pre-requisite. They reported that after getting familiar with different subcategories of the scoring matrix and experiencing self- and peer-assessment practices, they could more profoundly learn how to write, how to organize, and what points to consider when writing an essay. Likewise, they gradually learned how to professionally judge and comment on an academic five-paragraph essay. Some students emphasized that the experience of assessing essays has caused them to be more critical and mindful when writing their own essays, paying more attention to those areas stated in the evaluation criteria.

Most of the students found it more difficult to be critical in peer-assessment than in self-assessment because they were not familiar with the writing style, ideas, and attitudes of their peers towards the subject that they had developed.

They typically noticed that being critical about the content, organization, and syntax was the most demanding facet. However, some others found it more challenging to do self-assessment as it was more difficult for them to find their own writing errors.

Concluding remarks

As mentioned in the previous sections, the main purpose of the present research was to bring LOA into an essay writing course in order to investigate and compare different types of alternative assessment (i.e. self-, peer-, and teacher-assessments) and to see if there is any significant difference among them. The next aim was to probe into the experiences and perceptions of the learners towards such experience. The results showed that more knowledgeable students could score more realistically to both themselves and their peers, while less knowledgeable ones overestimated the writing capabilities, maybe because they were not capable of detecting the writing problems and distinguishing them from accurately organized sections. The positive point about this research was that it could help the participants put the scoring criteria into practice and get familiar with the matrix based on which their essays are scored by their teachers. They could learn how to have a more realistic look on a piece of writing and what items to consider in writing an appropriate essay. In this research, the role of students' attitudes towards LOA was not ignored, as the researchers believed that the scores given by the participants could not be reliable unless they involve the marking process with positive attitudes. Taking the role of students' attitudes into their scoring practices played a crucial role in crystallizing the significance of the study as almost all participants found their experience both useful and motivating.

References

Ahmadpour, L., & Yousefi, M. H. (2016). Group collaboration, scaffolding instruction, and peer assessment of Iranian EFL learners oral tasks. *The Journal of Applied Linguistics and Applied Literature: Dynamics and Advances (JALDA), 4*(1), 31-44.

Alderson, J. C., & Banerjee, J. (2001). Language testing and assessment (Part I). *Language Teaching, 34*(4), 213-236. doi:10.1017/s0261444800014464

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into Practice, 48,* 12-19. doi:10.1080/00405840802577544

Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment and Evaluation in Higher Education, 32*(2), 159-181. doi:10.1080/02602930600801928

Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that we count counts. *Language Testing 17*(1), 1-42. doi:10.1177/026553220001700101

Bell, J. (1991). Using peer responses in ESL writing classes. *TESL Canada Journal, 8*, 65-71. doi:10.18806/tesl.v8i2.589

Biggs, J. (1999). *Teaching for quality learning at university*. Buckingham: Open University Press.

Birdsong, T., & Sharplin, W. (1986). Peer evaluation enhances students' critical judgment. *Highway One, 9*, 23-28.

Birenbaum, M., & Dochy, F. (1996). *Alternatives in assessment of achievements, learning processes and prior knowledge*. Boston, MA: Kluwer Academic Publishers.

Blatchford, P. (1997). Students' self-assessment of academic attainment: Accuracy and stability from 7 to 16 years and influence of domain and social comparison group. *Educational Psychology, 17*(3), 345-360. doi:10.1080/0144341970170308

Boroditsky, L. (2001). Does language shape thought? Mandarian and English speakers' conception if time. *Cognitive Psychology, 43*, 1-22. doi:10.1006/cogp.2001.0748

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

Boud, D., & Falchikov, N. (2007). *Rethinking assessment in higher education.* London: Kogan Page.

Brown, D. (2004). *Language assessment: Principles and classroom practice*. New York: Longman.

Brown, G., Bull, J., & Pendlebury, M. (1997). *Assessing students' learning in higher education.* London: Routledge.

Brown, H. D. (2007). *Principles of language learning and teaching* (5th ed.). NY: Pearson Education.

Brown, J. D. (1991). *Breaking the language barrier.* Yarmouth, ME: Intercultural Press.

Brown, J. D., & Hudson, T. (2002).*Criterion-referenced language testing.* Cambridge: Cambridge University Press.

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

Cho, K., & MacArthur, C. (2010).Student revision with peer and expert reviewing. *Learning and Instruction, 20*, 328–338. doi:10.1016/j.learninstruc.2009.08.006

Cho, Y., & Cho, K. (2011). Peer reviewers learn from giving comments. *Instructional Science, 39*, 629-643. doi:10.1007/s11251-010-9146-1

Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second language writing.* Cambridge: Cambridge University Press. doi:10.1017/CBO9781139524599

Cyboran, V. (2006). Self-assessment: Grading or knowing? *Academic Exchange Quarterly, 10*(3), 183-186.

Davies, P. (2006). Peer assessment: Judging the quality of students' work by comments rather than marks. *Innovations in Education and Teaching International, 43*(1), 69-82. doi:10.1080/14703290500467566

Devenny, R. (1989). How ESL teachers and peers evaluate and respond to student writing. *RELC Journal, 20*, 77-90. doi:10.1177/003368828902000106

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*(3), 331-51. doi:10.1080/03075079912331379935

Earl, S.E. (1986). Staff and peer assessment: Measuring an individual's contribution group performance. *Assessment and Evaluation in Higher Education, 11*, 60-69. doi:10.1080/0260293860110105

Entwhistle, N.J. (1987). A model of the teaching process. In J. T. E. Richardson, M. W. Eysenck, & D. W. Piper (Eds.), *Student learning: Research in education and cognitive psychology.* Milton Keynes, UK: Open University Press.

Entwhistle, N.J. (1993). *Recent research on student learning and the learning environment.* Paper presented at the 'New Developments in Learning' Conference, Napier University, Edinburgh.

Eva, K. W. & Regehr, G. (2005). Self-assessment in the health professions: A reformulation and research agenda. *Academic Medicine, 80*(10), 546-554. doi:10.1080/0260293860110105

Freeman, M. (1995). Peer-assessment by groups of group work. *Assessment and Evaluation in Higher Education, 20*(3), 289-300. doi:10.1080/0260293950200305

Gibbs, G. (1992). *Improving the quality of student learning.* Bristol: Technical and Educational Services.

Habeshaw, S. Gibbs, G., & Habeshaw, T. (1995). *53 Interesting ways to assess your students.* Bristol: Technical and Educational Services.

Hamayan, E. V. (1995). Approaches to alternative assessment. *Annual Review of Applied Linguistics, 15,* 212-226. doi:10.1017/s0267190500002695

Hart, D. (1994). *Authentic assessment: A handbook for educators.* New York: Addison-Wesley.

Huerta-Macias, A. (1995). Alternative assessment: Responses to commonly asked questions. *TESOL Journal, 5*(1), 8-11. doi:10.1017/cbo9780511667190.048

Jafarpur, A. (1991). Can naive EFL learners estimate their own proficiency? *Evaluation and Research in Education, 5*, 145-57.

Kaplan, R. B. (1972). Contrastive rhetoric and second language learning: Notes toward a theory of contrastive rhetoric. In A. C. Purves (Ed.), *Writing across languages and cultures: Issues in contrastive rhetoric* (pp. 257-304). Newbury Park, CA: Sage.

Kohonen, V. (1997). Authentic assessment as an integration of language learning, teaching, evaluation and the teacher's professional growth. In A. Huhta, V. Kohonen, L. Kurki-Suonio, & S. Luoma (Eds.), *Current development and alternatives in language assessment* (pp. 6-15). Proceedings of LTRC 1996. University of Jyväskylä.

Langan, A. M., & Wheater, P. (2003). Can students assess students effectively? Some insights into peer-assessment. *Learning and Teaching in Action, 2*(1), Retrieved from http://www.celt.mmu.ac.uk/ltia/issue4/langanwheater.shtml

Lay, N. (1982). Composing processes of adult ESL learners. *TESOL Quarterly, 16(*3), 406.

LeBlanc, R., & Painchaud, G. (1985). Self-assessment as a second language placement instrument. *TESOL Quarterly, 19*(4), 673-687. doi:10.2307/3586670

Lejk, M. & Wyvill, M. (2001). The effect of the inclusion of self-assessment with peer-assessment of contributions to a group project: A quantitative study of secret and agreed assessments. *Assessment and Evaluation in Higher Education, 26*(6), 551-561. doi:10.1080/02602930120093887

Lew, M. D. N, Alwis, W. A. M, & Schmidt, H. G. (2010). Accuracy of students' self-assessment and their beliefs about its utility. *Assessment & Evaluation in Higher Education, 35*(2), 135-156. doi:10.1080/02602930802687737

Lindblom-Ylänne, S., & Lonka, K. (2001). Students' perceptions of assessment practices in a traditional medical curriculum. *Advances in Health Sciences Education, 6*(2), 121-40. doi:10.1023/a:1011422517238

Lindblom-Ylänne, S., Pihlajamäki, H., & Kotkas, T. (2006). Self-, peer- and teacher-assessment of student essays. *Active Learning in Higher Education, 7*(1), 51-62. doi:10.1177/1469787406061148

Magin, D., & Helmore, P. (2001). Peer and teacher assessments of oral presentations: How reliable are they? *Studies in Higher Education, 26*, 287-298. doi:10.1080/03075070120076264

Mangelsdorf, K. (1992). Peer reviews in the ESL composition classroom: What do students think? *ELT Journal, 46*, 158-72. doi:10.1093/elt/46.3.274

Martin-Kniep, G. O. (2000). Standards, feedback, and diversified assessment: Addressing equity issues at the classroom level. *Reading & Writing Quarterly, 16,* 239-256. doi:10.1080/105735600406733

Mendonca, C., & Johnson, K. (1994). Peer review negotiations: Revision activities in ESL writing instruction. *TESOL Quarterly, 28*, 745-69. doi:10.2307/3587558

Mousavi, A. (2011). *An encyclopedic dictionary of language testing* (5th ed.). Tehran: Rahnama Press.

Nilson, L. B. (2003). Improving student peer feedback. *College Teaching, 51*(1), 34–38. doi:10.1080/87567550309596408

O'Malley, M., & Valdez Pierce, L. (1996). *Authentic assessment for English language learners.* New York: Addison-Wesley Publishing Company.

Oldfield, K., & MacAlpine, M. (1995). Peer and self-assessment at tertiary level an experiential report. *Assessment and Evaluation in Higher Education, 20*, 125-32. doi:10.1080/0260293950200113

Orsmond, P., Merry, S., & Reiling, K. (1996). The importance of marking criteria in peer assessment. *Assessment and Evaluation in Higher Education, 21*(3), 239-49. doi:10.1080/0260293960210304

Orsmond, P., Merry, S., & Reiling, K. (1997). A study in self-assessment: Tutor and students' perceptions of performance criteria. *Assessment and Evaluation in Higher Education, 22*(4), 357-69. doi:10.1080/0260293970220401

Orsmond, P., Merry, S., & Reiling, K. (2000). The use of student derived marking criteria in peer and self-assessment. *Assessment and Evaluation in Higher Education, 25*(1), 23-38. doi:10.1080/02602930050025006

Orsmond, P., Merry, S., & Reiling, K. (2002). The use of exemplars and formative feedback when using student derived marking criteria in peer- and self-assessment, *Assessment and Evaluation in Higher Education, 27*(4), 309-23. doi:10.1080/0260293022000001337

Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing, 6*(1), 1-13. doi:10.1177/026553228900600103

Patri, M. (2002).The influence of peer feedback on self- and peer-assessment of oral skills. *Language Testing, 19*(2), 109-131. doi:10.1191/0265532202lt224oa

Peng, J. (2010). Peer assessment in an EFL context: Attitudes and correlations. In M. T. Prior et al. (Eds.), *Selected proceedings of the 2008 Second Language Research Forum* (pp. 89-107). Somerville, MA: Cascadilla Proceedings Project.

Race, P. (1999). *2000 Tips for lecturers*. London: Kogan Page.

Rolheiser, C., & Ross, J. A. (2000). Student self-evaluation-What do we know? *Orbit, 30*(4), 33-36.

Ross, J. A. (2006). The reliability, validity and utility of self-assessment. *Practical Assessment, Research & Evaluation, 11*(10), 1-13. doi:10.1007/0-306-48125-1_4

Segers, M., & Dochy, F. (2001). New assessment forms in problem-based learning: The value added of the students' perspective, *Studies in Higher Education, 26*(3), 327-43. doi:10.1080/03075070120076291

Sivan, A. (2000). The implementation of peer assessment: An action research approach. *Assessment in education: Principles, policy & practice, 7*(2), 193-213. doi:10.1080/713613328

Smith, K. (1999). Language testing: Alternative methods. In B. Spolsky (Ed.). *Concise encyclopedia of educational linguistics*, (pp. 703-706). Amsterdam: Elsevier.

Soodak, L. C. (2000). Performance assessment: Exploring issues of equity and fairness. *Reading and Writing Quarterly, 16,* 175-178. doi:10.1080/105735600406706

Spiller, D. (2012). *Assessment matters: Self-assessment and peer assessment. Teaching Development Unit, the University of Waikato.* Hamilton, New Zealand.

Stefani, A.J. (1994). Self, peer and group assessment procedures. In I. Sneddon & J. Kramer (Eds.), *An enterprising curriculum: Teaching innovations in Higher education* (pp. 24-46). Belfast: HMSO.

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249-276. doi:10.2307/1170598

Topping, K. (2003). Self- and peer-assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.) *Optimizing 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*(1), 20-27. doi:10.1080/00405840802577569

Willey, K. & Garner, A. (2010). Investigating the capacity of self & peer assessment activities to engage students & promote learning. *European Journal of Engineering Education, 35*(4), 429-443. doi:10.1080/03043797.2010.490577

Williams, E. (1992). Student attitudes towards approaches to learning and assessment. *Assessment & Evaluation in Higher Education, 17*(1), 45–58. doi:10.1080/0260293920170105

Worthen, B. R. (1993). Critical issues that will determine the future of alternative assessment. *Phi Delta Kappan, 74*(6), 444-456. doi:10.1177/003172170808900811

Zoller, U., & Ben-Chaim, D. (1997, August). *Student self-assessment in HOCS science examinations: Is it compatible with that of teachers?* Paper presented at the 7th EARLI conference, Athens, Greece.

Author Biography

Gholam-Reza Abbasian, born in Ahar, East Azerbaijan, is an Assistant Professor of TEFL at Imam Ali and IA (South Tehran) universities, and has presented some papers at national and international conferences. He is the author and translator of about 15 books, and publisher of scholarly articles. Dr. Abbasian offers psycholinguistics, language testing, and research methods at MA and PhD levels. Nominated as top scholar and teacher for seven consecutive years, he is the internal manager of JOMM, reviewer of Sage, FLA, and GJER journals and a member of editorial board of JSSIR.

Maryam Khezrinejad is currently a Ph.D. candidate of TEFL at Islamic Azad University, Science and Research Branch, Tehran, Iran. She is a faculty member and the head of language department at Islamshahr Azad University. She has taught several English courses to BA students of English for over a decade. She has authored a number of books and has published international articles.

Mojtaba Teimourtash is currently a Ph.D. candidate of TEFL at Islamic Azad University, Science and Research Branch, Tehran, Iran. He is an English instructor.