Self- vs. Peer-editing: One Step Forward from Assessment to Building EFL Students’ Writing Skill

Fatemeh Behjat*
Department of Foreign Languages
Ph.D candidate in TEFL
Islamic Azad University, Abadeh Branch
Abadeh, Iran
Email: fb-304@yahoo.com

Mortaza Yamini
Department of Foreign Languages
Assistant Professor
Shiraz University
Shiraz, Iran

Abstract

A number of pedagogical arguments support self- and peer-assessment in language classrooms to improve the quality of learning by students’ involvement in the final judgments of one’s work. Considering editing as a method of communicative testing of writing, this study aimed at finding out whether self- and peer-editing of assignments could result in the improvement in Iranian EFL students’ writing skill. For this purpose, 90 sophomore English students at Shiraz Islamic Azad University were selected and assigned a topic to write about as the pre-test. The participants were then divided into two groups and trained for assessing writing. While in the first group, the participants rated their own writings, in the second, they were asked to correct their peers’ writings. The participants then wrote a paragraph on a topic as the post-test. The result of a two-way ANOVA for the comparison between the participants’ performance in the pre- and post-tests revealed an improvement in the students’ writing ability, and there was no significant difference between the performances of the two groups implying that self- and peer-assessment can help the EFL learners equally to improve their writing.

Keywords: peer-editing, self-editing, communicative writing, indirect method for assessing writing.

Received: April, 2011; Accepted: October, 2011

*Corresponding author
1. Introduction

Peer- and self-assessment, as two methods of learner assessment have gained momentum in recent years. Peer- and self-assessment can be defined as a process through which individual students are given an opportunity to assess their or their peers’ writing abilities (Gennip, Nanine, Segers, and Tillema, 2009). Course evaluations provide teachers with feedback that generally focus on what the teacher does. But with student self-assessment it becomes possible to focus more directly on the student. This kind of assessment includes the evaluation of not only the outcomes but also the process of learning. One of the skills for which peer- and self-assessment can best be applied is writing. Johnson-Bogart’s (2000) study revealed that students’ writings after peer and self assessment showed that they learned from each other and wanted to contribute to each other; they wanted to be useful. And what he found was that learning from one another and learning from their own mistakes were mutually reinforcing. According to Weir (1990), one of the indirect methods for assessing learners’ writing ability is editing in which the student is given a text and is asked to rewrite the passage making all the necessary corrections.

Learning to write at an acceptable academic level cannot be isolated from learning a foreign language as writing communicatively is one of the skills in language learning. Yet, as learning other language skills, the acquisition of academic writing is a long-term endeavor. Therefore, in many curricula, teachers search for proper methods to provide more support to students in developing their writing competence. Peer and self assessment are activities in which students consider the quality of their
Self- vs. Peer-editing: One Step Forward...

fellow students’ work as well as that of themselves and in which the assessment is a formative one.

Self- and peer-assessment have always been used side by side as two tools to improve writing ability. The problem which is raised here is which of them can be more fruitful than the other. Following Berg, Wilfried and Pilot (2003), who supported the influential role of peer and self assessment in the development of learners’ writing skill, the purpose of the present study was to find out whether self- and peer-editing could improve the writing ability of Iranian EFL learners. This study also aimed at revealing which of the editing tasks, peer or self, resulted in the improvement of language learners’ writing performance better, and whether male and female students practicing peer- and self-editing improve their writing ability at the same rate.

2. Literature Review

The reason why it is desirable and sensible to have students involved and central to the assessment process is illustrated by Boud and Falchikov (1989). It is stated that teachers have limited access to their students’ knowledge and in many ways students have greater insights into their own achievements. The number of empirical research studies on peer and self involvement in language classrooms is growing.

The defining characteristic of self-assessment, according to Orsmond (2004), is the involvement of students in identifying standards and criteria to apply to their own work and making judgments about the extent to which they have met these criteria and standards. Peer assessment has been defined (Topping, Smith, Swanson, and Elliot,
as an arrangement for peers to consider the level, value, quality or successfulness of the products or outcomes of learning of others of similar status. From these definitions, it becomes apparent that self- and peer-assessments are sources of assessment that may be used within a framework of working on any language skill (Brown, Bull, and Pendlebury, 1997).

Peer involvement in assessment has a great potential for learning. Even though evidence from language classrooms is limited, peer assessment has proved to have a number of benefits: it has a strong relationship with the instructor’s ratings in different settings (Falchikov and Goldfinch, 2000); it encourages reflective learning through observing others’ performances and becoming aware of performance criteria (Topping, 1998); and it generates positive reactions for students, and the students can develop a sense of shared responsibility toward others’ work (Saito and Fujita, 2004; Somervell, 1993).

Peer and self assessment can guarantee bringing benefits to students’ learning as they are capable of implementing the assessment. A number of studies have shown that training can help learners to rate themselves and their peers better than untrained ones in such a way that the trained group’s ratings can correlate with instructors’ ratings. Shohamy, Gordon, and Kraemer, (1992) found positive effects of training and stated that trained raters are more reliable than untrained ones. Jafarpur and Yamini (1995) showed that training with self-assessment and peer-rating questionnaires improved language learners’ skill to estimate their own and their peers’ language ability. McGroarty and Zhu (1997) found that trained groups spent more time on revision and produced more comments.
Self- vs. Peer-editing: One Step Forward...

than whose who were untrained. In Weigle’s study (1998), novice raters were more strict and inconsistent than experienced raters.

Regarding the difference in the performance of male and female learners in self and peer assessment, Wheater, Langan, and Dunleavy (2005) stated that while female peers are more precise in their ratings male students are more lenient and give better marks to their classmates in assessing their written work.

Research on assessing and developing students’ writing emphasizes how complex writing is and how it involves deep aspects of students’ understandings of themselves and their ability and willingness to participate in the academic community (Crème, 2000, Ivanic, 1998). Cowan (2004), in his study on the role of reflective self and peer assessment in writing evaluation, concluded that although initially fearful, most students found the experience of reading a peer’s work helpful and enjoyable, and that self assessment makes students become more confident and autonomous in writing. He stated that peer exercise should be paired with a self-assessment exercise. Self and peer-assessment are often combined or considered together. Peer assessment can help self-assessment. By judging the work of others and themselves, students gain insights into their own performance. Peer and self-assessment help students develop the ability to make judgments, which is a necessary skill for learning (Graham and Rachel, 1995). According to Lillis (2001), self and peer assessment in writing are socially situated activities, involving issues of building a better social identity and social relations. Meldrum (2002) argued that while self-assessment provides students with more autonomy to judge their own work, more is known
about the students in terms of how they view themselves.

Orsmond (2004) compared self- and peer-assessment with other assessments and concluded that the first two are student-centered, have transparent criteria, encourage a deep approach to learning, allow students to construct their learning, encourage discussions between students and tutors, give opportunity to revise weak areas of learning, increase the students’ confidence, increase learning quality of the output, and offer authentic learning tasks. Senges (2008) refers to peer-assessment as a ‘collabowriting’ in which students involve in each others’ writing practices and improve their writing ability.

As for communicative tests, Mousavi (1999) defined it as a test which requires candidates to perform communicatively like that of real-life situations. A communicative test has to meet some criteria. It has to test for grammar as well as the content and organization. Communicative tests have task dependency. The content of communicative tests is motivating and interactive. McNamara (2000) stated that communicative tests ultimately come to have two features: they require assessment to be carried out when the candidate is engaged in an act of communication, whether productive or receptive, and they pay attention to the social roles candidates are likely to assume in real world settings. Hughes (2003) mentioned that in order to test writing we need to set writing tasks that are representative of the tasks we expect the students to be able to perform, and samples of writing should be scored holistically and analytically. To Elder, Brown, Grove, Hill, Iwashita, Lumley, McNamara, and O’Loughlin (2000), communication involves both identifying separate features as well as the relations between them.
Communicative testing must be devoted not only to what the learner knows about the second language and how to use it but also to what extent the learner is able to actually demonstrate this knowledge in a meaningful communicative performance (Fulcher and Davidson, 2007). Weir (1990) suggested two different approaches for assessing writing ability communicatively. They include direct and indirect methods. Indirect methods like editing task enjoy the advantage of having a good washback effect as students are taught and encouraged to edit their written work more carefully. Students are often unfamiliar with marking criteria. Hence, they need to be clearly introduced to them at the beginning of their course of study. Therefore, Weir (1999) offered appropriate criteria for assessing written production: the Test in English for Educational Purposes (TEEP). In this scale, the criteria of relevance and adequacy, organization, cohesion, grammatical accuracy, spelling, and punctuation were seen as the most suitable factors for assessing writing tasks (see the Appendix for more details). To him, in order to resolve the issue of criteria, both the teacher and the students should be able to clarify the concepts of the criteria.

It is worth mentioning here that different types of criteria can be used to generate different forms of judgments. Miller (2003), for example, considered the implementation of criteria within a self- and peer-assessment context. Miller, looking at oral presentation over two consecutive years, wanted to change assessment marking from looking at a few global components of performance to very discrete ones. Anyway, making judgment is ultimately what assessment is about. Having explicit and unambiguous criteria helps this process, but regarding students’ writing assignments, it is still a challenge for them to
take their own or their peers’ work home and make judgments about it. While peers may be unwilling to make formal assessment of their peers, they may be more positive when students have to give specific feedback of a descriptive nature for the benefit of their peers and no grading has taken place (Boud and Falchikov, 1989).

Berg et al. (2003) did a study on the contribution of peer assessment to the acquisition of communicative writing skills by university students. Additionally, they aimed at establishing an optimal model of peer assessment. Aspects that were considered included the implementation of peer and self assessment by students and teachers in practicing writing, the components of peer feedback, the interaction between students during oral peer feedback, students’ achievement and students and teachers’ evaluation of peer assessment. The results showed that peer assessment could improve students’ writings to a considerable extent.

Following Berg et al. (2003), Orsmond (2004), and Senges (2008) on the effectiveness of peer and self assessment, this study aims at finding out which method, peer- or self-editing, can improve writing performance of Iranian EFL writers’ better than the other. Furthermore, another objective of the present study is to see which method can work better for male and female learners.

### 2.1 Research questions and hypotheses

Following Berg et al. (2003), the following research questions were posed in the present study:

Q1- Do self- and peer- editing improve the writing performance of Iranian EFL learners?
Q2- Do both self- and peer-editing groups improve their writing ability in the same way?

Q3- Do male and female students practicing peer- and self-editing improve their writing ability in the same way?

Correspondingly, three null hypotheses were formed based on the above-mentioned research questions:

NH1- Self- and peer-editing do not improve the writing ability of Iranian EFL learners.

NH2- Self-editing does not improve the writing performance of Iranian EFL students as peer-editing does.

NH3- Male students do not perform well in their writing performance as the female students do using self- and peer-editing.

3. Method

3.1 Participants
The participants of the present study were 14 male and 76 female sophomore students majoring in English at Shiraz Islamic Azad University. As the number of female EFL students is usually about three times more than male students, the researcher could not have approximately the same number of male and females in this study. They had all passed their Grammar and Writing (1) and (2) courses and were taking Advanced Writing course; therefore, they were assumed to have the same background knowledge of English grammar and writing. After the participants showed the same level of writing ability through a pre-test, they were divided into two groups of 45 students. All the
participants were at the average age range of 21 to 26 with a Standard Deviation of 1.57.

3.2 Instrumentation
All the participants were asked to write one paragraph on assigned topics before the instruction, two weeks after the beginning of the instruction, and the end of instruction as their pre- and post-tests of writing ability. Later the researcher used the first two writing test results to compare the consistency of the scores in rating the students’ writings. In other words, the first two writing tests were taken to check intra-rater reliability based on test-retest method.

During the treatment, the participants were asked to write a paragraph on an assigned topic each session; then they were rated and corrected either by themselves or by their classmates after two sessions of the delivery of papers to the instructor.

The TEEP Attribute Writing Scale developed by Weir (1990) was used as the criteria to help the participants rate their own papers or those of their peers. The scale had seven components including relevance and adequacy of content, organization, cohesion, appropriate vocabulary, grammar, and mechanical accuracy (punctuation and spelling). The specifications for the rating scale are given in the Appendix.

3.3 Procedure
In order to see if the participants were truly randomly chosen from among all available EFL students, and thus they are homogeneous, a topic was assigned to them at the beginning of the instruction to write a
paragraph about. As only a valid and reliable assessment process should be used to determine whether learning has occurred, the content of the present writing test was consulted with a number of colleagues for content validity. It was, then, revealed that the test enjoyed a high level of content validity. To ensure the reliability of the scores, the students were asked to write another paragraph two weeks later, and then the scores on both tests were correlated for the intra-rater reliability. As the result turned out to be .82, it was concluded that the scores of the tests were reliable, too.

In the first session of the instruction period, the instructor trained all the participants on how to rate one’s writing. For this purpose, Weir’s (1990) TEEP rating scale was used. To apply the criteria an attempt is made in this scale to construct an analytic marking scheme in which each of the criteria is sub-divided into four levels of 0-3 (see the Appendix). A level three corresponds to a base line of minimal competence. At this level it is felt that a student is likely to have very few problems in coping with the writing task. At level two a limited number of problems arise in relation to the criterion and remedial help would be advisable. A level one would indicate that a lot of help is necessary with respect to this particular criterion. A level zero indicates almost total incompetence in respect of the criterion in question.

From the second session on, for a period of two months, the participants were asked to write paragraphs on the assigned topics. The papers were collected and then the papers were brought to the classroom by the teacher. For the first group, the papers were given back to the writers of the papers to be rated and corrected. This group had the self-editing
task. In the second group, however, the papers were handed in to the students, but this time, they were exchanged among the peers to be rated based on Weir’s (1990) TEET scale. The students were asked to make the necessary corrections and give back the papers to their owners.

At the end of the course, the participants again took two paragraph writing tests with the same topic for both groups. As the intra-rater reliability for both post-tests of writing was at a good level, one set of scores were considered as the post test. The scores on the pre and post test, then, were entered into the computer for analysis.

4. Data Analysis

Before the treatment started, an independent samples \( t \)-test was run on the pre-test scores of the two groups to make sure they were homogeneous and truly randomly selected. The results are shown in Table 1.

<table>
<thead>
<tr>
<th>( t )</th>
<th>df</th>
<th>sig</th>
<th>Mean dif</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>.49</td>
<td>88</td>
<td>.627</td>
<td>.2667</td>
<td>-.821 - 1.355</td>
</tr>
</tbody>
</table>

As Table 1 suggests, the value of observed \( t \) (\( t = .49 \)) is small and not statistically significant (sig. = .627); therefore, it can be concluded that both groups were at the same level as far as their writing was.
After the instruction, both groups took part in another writing test as their post-test. This time, too, the participants were assigned a topic for a paragraph. The papers were corrected and the scores were entered into the computer. A paired t-test was run to see the difference in students’ performance between the pretest and post-test. The results are shown in Table 2.

**Table 2. t-test between pre-test and post-test**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>t</th>
<th>df</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>pretest-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>posttest (peer)</td>
<td>-3.044</td>
<td>18.</td>
<td>4</td>
<td>.00</td>
</tr>
<tr>
<td>pretest-</td>
<td>-1.467</td>
<td>14.</td>
<td>4</td>
<td>.00</td>
</tr>
<tr>
<td>posttest (self)</td>
<td></td>
<td>9</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

As can be seen from the table, both groups have a significant mean difference. This means that the instruction has been effective. Therefore, the first null hypothesis stating that there would be no improvement in writing ability of EFL learners is rejected.

In order to answer the second and third research questions, a two-way ANOVA was run on gain score as a dependent variable and group and sex as independent variables. The results are presented in Tables 3 and 4.
Table 3. Descriptive statistics on gain score

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>14</td>
<td>1.0</td>
<td>5.0</td>
<td>2.85</td>
<td>1.61</td>
</tr>
<tr>
<td>Female</td>
<td>76</td>
<td>0.00</td>
<td>5.0</td>
<td>2.14</td>
<td>1.07</td>
</tr>
<tr>
<td>Self</td>
<td>45</td>
<td>0.00</td>
<td>3.0</td>
<td>1.46</td>
<td>.660</td>
</tr>
<tr>
<td>Peer</td>
<td>45</td>
<td>1.0</td>
<td>5.0</td>
<td>3.04</td>
<td>1.08</td>
</tr>
</tbody>
</table>

Table 4. ANOVA results on gain score

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>37.22</td>
<td>1</td>
<td>37.22</td>
<td>47.9</td>
<td>.00</td>
</tr>
<tr>
<td>Sex</td>
<td>1.357</td>
<td>1</td>
<td>1.357</td>
<td>1.74</td>
<td>.18</td>
</tr>
<tr>
<td>Group * Sex</td>
<td>1.998</td>
<td>1</td>
<td>1.998</td>
<td>2.57</td>
<td>.11</td>
</tr>
</tbody>
</table>

78
Self- vs. Peer-editing: One Step Forward...

As Table 4 shows, the difference between groups is significant (F=47.977, sig. = .000). This means that the two groups did not develop their writing ability in the same way. When the means reported in Table 3 are considered, it becomes clear that the peer group (mean = 3.044) outperformed the self group (mean = 1.4667). In other words, the students benefited from peer-assessment more than self-assessment. The second null hypothesis is rejected.

The difference between males and females was not statistically significant (F = 1.749, sig. = .189), nor was the interaction between group and sex (F = 2.576, sig. = .112). The third null hypothesis is retained.

According to the above-mentioned results, while both peer and self editing can be considered as useful practices to do to foster writing ability of EFL students, peer editing can lead to a better writing performance compared to self editing. It means that if students are encouraged to read each others’ writings and make corrections on each others’ work, they can enhance their writing ability better compared to the time the students are assumed to read their own writings and try to edit them. Besides, the results of the study indicated that as far as learners’ gender is concerned, there is no significant difference in the performance of male and female learners in both self and peer groups.

5. Conclusion

This study aimed at finding out whether self- and peer-editing, as two methods of assessing communicative writing ability, could help the EFL learners’ writing ability or not. The results revealed that these two
methods of assessing can serve the purpose of improving writing ability of EFL learners to a considerable extent. Thus, the results of this study are in support of Orsmond(2004) regarding the positive effect of self- and peer-assessment on the improvement of language abilities of foreign language learners. The findings also indicated that students benefited from peer-assessment more than they benefited from self-assessment. Additionally, it was revealed that there was no significant difference between male and female students in their performance in writing.

6. Pedagogical Implications
The results of the present study can serve not only language teachers but also language testers and curriculum and syllabus designers. Teachers can determine a scale for assessing writing in the first session of their writing class and ask the students to edit their own writings and those of their peers in order to help them not only to improve their writing skill but also to strengthen their cooperative learning, self-confidence and social relations.

Language testers can find the amount of relationship between their own ratings with those of students, and thereby see how well the students can judge their own performance and that of their peers in writing. They can also use self- and peer-assessment in testing their students’ speaking skill.

Syllabus designers and curriculum developers can also include self- and peer-editing tasks as a part of an EFL program to help the learners foster their language skills. These methods of assessment cannot only be used in writing courses but they can be applied in oral presentation classes to
Self- vs. Peer-editing: One Step Forward...

assess students’ language proficiency.

The authors

Fatemeh Behjat is at present an ABD in TEFL at Islamic Azad University, Shiraz Branch and has been a faculty member at Islamic Azad University, Abadeh Branch since 2000. She taught English at the Iran Language Institute for eight years and is teaching at Zand Institute of Higher Education. She has so far presented papers at International conferences in Iran and Abroad and published articles in different journals. Her main area of interest are language teaching and learning.

Dr. Morteza Yamini is an assistant professor who taught for 35 years at Shiraz University before he retired in 2008 by request. He is still pursuing his teaching profession, offering some graduate courses at Shiraz University and undergraduate courses at Zand Institute of Higher Education. He is also an active member of the Center of Excellence in L2 Reading and Writing, Shiraz University.

References


F. Behjat and M. Yamini


Self- vs. Peer-editing: One Step Forward...


**Appendix: TEEP Writing Scale (Weir, 1990)**

**A. Relevance and adequacy of content**

0. The answer bears almost no relation to the task. Totally inadequate answer.

1. Answer of limited relevance to the task. Possibly major gaps in treatment of topic and pointless repetition.

2. For the most part answers the task, though there may be some gaps or
redundant information.

3. Relevant and adequate answer to the task.

**B. Organization**

0. No apparent organization of content.

1. Very little organization of content. Underlying structure not sufficiently controlled.

2. Some organizational skills in evidence, but not adequately controlled.

3. Overall shape and internal pattern clear. Organizational skills adequately controlled.

**C. Cohesion**

0. Cohesion almost totally absent. Writing so fragmentary that comprehension of the intended communication is virtually impossible.

1. Unsatisfactory cohesion may cause difficulty in comprehension of most of the intended communication.

2. For the most part satisfactory cohesion though occasional deficiencies may mean that certain parts of the communication are not always effective.

3. Satisfactory use of cohesion resulting in effective communication.

**D. Appropriateness of vocabulary**

0. Vocabulary inadequate even for the most basic parts of the intended communication.

1. Frequent inadequacies in vocabulary for the task. Perhaps frequent lexical inappropriateness or repetition.
Self- vs. Peer-editing: One Step Forward...

2. Some inadequacies in vocabulary for the task. Perhaps some lexical inappropriateness or circumlocution

3. Almost no inadequacies in vocabulary for the task. Only rare inappropriateness.

E. Grammar

0. Almost all grammatical patterns inaccurate.

1. Frequent grammatical inaccuracies.

2. Some grammatical inaccuracies.

3. Almost no grammatical inaccuracies.

F. Mechanical accuracy 1 (punctuation)

0. Ignorance of conventions of punctuation.

1. Low standard of accuracy in punctuation.

2. Some inaccuracies in punctuation.

3. Almost no inaccuracies in punctuation.

G. Mechanical accuracy 2 (spelling)

0. Almost all spelling inaccurate.

1. Low standard of accuracy in spelling

2. Some inaccuracies in spelling.

3. Almost no inaccuracies in spelling