

The Impact of Self-Assessment on Iranian EFL Learners' Writing Skill

Mohammad Reza Javaherbakhsh

Science and Research Branch of Islamic Azad University, Tehran, Iran

E-mail: r_javaheribakhsh@yahoo.com

Abstract

Self-assessment has gained much attention in recent years owing to graving emphasis on learner independence and learner autonomy and self-assessment has significant pedagogic value (Mrudula, 2002). The present study was an attempt to investigate whether self-assessment impact Iranian EFL learners' writing skill. To fulfil the purpose of the study, 73 participants from 105 students of advanced level of English in Zabansara and Marefat Institutions were chosen by means of a TOEFL test. Two groups; namely, experimental and control groups, were formed. A writing pre-test was administered to measure their writing ability. The subjects in the experimental group enjoyed the treatment. After each writing activity, they assessed themselves with writing checklist. Three raters scored the activities by utilizing the analytical method of scoring. The teacher commented on the students' weaknesses at the bottom of each checklist. At the end of the study, all subjects took writing post-test. By utilizing an independent t-test, the researcher found out that the change in the experimental group was significant. Therefore, the null hypothesis of the study was rejected. The results of this research can be used by those involved in the field of language teaching.

Keywords: Self-assessment, Iranian language learners, Writing skill

1. Introduction

Students need to know what their abilities are, how much progress they are making and what they can (or can not yet) do with the skills they have acquired. Without such knowledge, it would not be easy for them to learn efficiently. From an educational viewpoint, knowing to what extent students can appraise their own performance is also important. If they can do it accurately enough, they do not have to depend completely on the opinions of teachers, and at the same time, they can make teachers aware of their individual needs. John Upshur (1975, cited in Heilenman, 1990) was one of the first who provided a rationale for the use of self-assessment in the measurement of the second language abilities. He pointed out that learners have access to the entire gamut of their success and failure in the use of the second language, whereas any test of actual language use can sample only a small proportion of that ability. Since then, there has been a growing interest in the use of self-report measures in the assessment of second language competencies.

In line with the theories of learner autonomy, self-assessment is currently playing an important role in language teaching. The procedure involves students in judging their own learning, particularly their achievement and learning outcomes. Many have argued that teachers should help students construct knowledge through active involvement in assessing their own learning performance, and that students are empowered by gaining ownership of their learning and life-long learning skills (Chen, 2008). Teachers should provide opportunities for students to assess their language level so as to help them focus on their own learning (Oskarsson, 1989). Hunt, Gow, and Barnes (1989) argue that without learner self-evaluation and self-assessment "there can be no real autonomy" (p. 207). Blue (1994) identifies some benefits of self-assessment such as encouraging greater effort, boosting self-confidence, and facilitating awareness of distinctions between competence and performance as well as self-awareness of learning strengths and weaknesses.

Self-assessment is used to prepare students for effectiveness and improvement in their lives. Self-assessment is thus linked with the goal of life-long learning and integrated into various subjects and domains. It has become not only a means to an end (autonomous life-long learning), but an end itself (a crucial component of autonomy). Through the process of self-assessment, students learn to discern patterns of strength and weakness that can help them become better learners. Self-assessment skills help students gradually develop a critical attitude toward learning throughout their lives and then achieve perfect autonomy.

Along with these theories of learning, the practice of self-assessment also reflects new thinking about classroom assessment: assessment for learning and evaluation. The assessment for learning approach shifts the focus from summative to formative assessment, from making judgments to prove that students have learned to provide feedback to indicate that they are making progress (Chen, 2008). Therefore, this study attempts to investigate whether learners' self-assessment of learning process influences the development of their writing skill.

Objectives of the Study

To fulfil the purpose of this study, the following research question is addressed:

Does self-assessment significantly impact Iranian EFL learners' writing skill?

Null Hypothesis: Self-assessment does not significantly influence Iranian EFL learners' writing skill.

2. Review of Literature

Erwin (1991) defines assessment as "the process of defining, analyzing, interpreting, and using information to increase students' learning and development" (p.15). Assessment is a systematic approach to collect information and make inferences about students' capability or the quality or success of a teaching course according to the different sources of the students' performance. There are various ways to do assessment. Such as test, interview, questionnaire, observation, etc. For instance, it is necessary to assess the comprehension ability of an immigrant student to understand if the student can follow a course of study in a school, or extra language teaching should be taken. Students might be tested both at the beginning and at the end of a course to assess the quality of the teaching on the course (Richards, & Schmidt, 2002). They may also be tested at the both sides of the teaching process to assess their level of ability. According to the sample that is taken, inferences are made about a person's achievements, potential, intelligence, aptitudes, attitudes, motivations, etc. and then the worth in the form of grades, recommendations are made (Mousavi, 1999).

Many educators have come to recognize that alternative assessments are an important means of gaining a dynamic picture of students' academic and linguistic development. "Alternative assessment refers to 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" (Hamayan, 1995, p.213). It is particularly useful with English as second language students because it employs strategies that task students to show what they can do. In contrast to traditional testing, in alternative-assessment students are evaluated on what they integrate and produce rather than on what they are able to recall and reproduce" (Huerta- Macias, 1995, p.9). Although there is no single definition of alternative assessment, the main goal is to "gather evidence about how students are approaching, processing, and completing real-life tasks in a particular domain" (Huerta-Macias, 1995, p.9). Alternative assessments generally meet the following criteria:

- Focus is on documenting individual student growth over time, rather than comparing students with one another.
- Emphasis is on students' strengths (what they know), rather than weaknesses (what they do not know).
- Consideration is given to the learning styles, language proficiencies, cultural and educational backgrounds, and grade levels of students.

Richards and Renandya (2002) also state that alternative assessment provides alternatives to traditional testing in that it a) does not intrude usual classroom activities; b) reflects the curriculum that is actually being implemented in the classroom; c) provides information on the strengths and weaknesses of each individual student; d) provides evidence that can be used to measure student progress; and e) is more multiculturally sensitive and free of norm, linguistic, and cultural biases found in traditional testing.

Moreover, Hamayan (1995) believes that some features of alternative assessment are as follows:

- Assessment is based on authentic tasks that demonstrate learners' ability to accomplish communication goals.
- Instructor and learners focus on communication, not on right and wrong answers.
- Learners help to set the criteria for successful completion of communication tasks.
- Learners have opportunities to assess themselves and their peer

Mousavi(1999) believes that self-assessment is the students 'own evaluation of their language ability, according to their capability to use the language in different situations. Self-assessment is a way to collect information about a person's knowledge in a language. Students' role shouldn't be restricted to the study of content chosen by others, through techniques selected by others. Researchers have shown that self-assessment develops students' motivation.

It is important for a learner to prepare for autonomy to be able to make some kind of judgment about accuracy of his performance. Because self-assessment emphasizes learning, the process rather than the results, or the product, it is seen, as one of the pillars of learner autonomy. "One of the fundamental elements of self-directed language learning is the opportunity given to learners to assess their own progress and thus help them to focus on their own learning" (Harris, 1997, p.12).

Dikel (2005) also argues that self-assessment is a process by which the student learns about himself, what he likes, what he doesn't like, and how he tends to react to certain situations. In the words of Rubin (1975), a good language learner monitors his own speech and the speech of others. That is, he is constantly attending to how well his speech is being received and whether his performance meets the standards he has learned.

In the last decade, with the increased attention to learner-centred curricula, needs analysis, and learner autonomy, the topic of self-assessment has become of particular interest in testing and evaluation (Blanche, 1988). Self-assessment is used by learners to evaluate and monitor their own level of knowledge, performance, and understanding and to get information about their learning (Dickinson, 1987). Therefore, self-assessment is what the students see from their own perspectives. Enabling students to self-monitor their own learning process is a way to help them develop knowledge through conscious control over that knowledge or to develop metacognitive awareness of knowledge and thought. Oskarsson (1989) thinks that self-assessment can promote learning because it gives learners training in evaluation which is important for autonomous learning. Learners need to be able to make reliable and valid judgments. The use of self-assessment in self-access learning is suggested because it enables the learners to reflect on their progress. This reflection enables the learners to take more control of their learning (Srimavin, & Daraswang, 2003).

3. Methodology

3.1 Participants

The participants of the present study were 73 students of advanced level in Merefat and Zabansara Language institutes in Tehran. Forty of them served as the experimental group and thirty-six as the control group. Ten students were eliminated because they did not take part in the post-test. All students were male, and the classes were intact. The participants of this study were selected from among a population of 105 students on the basis of the results of a standardized proficiency test. They were all native speakers of Persian and learners of English as a foreign language.

3.2 Instrumentation

The instruments used in this study include: a TOEFL test, a rating checklist, and two writing tests; i. e., pre and post-tests.

A proficiency TOEFL test was administered to measure the subjects' language knowledge. The TOEFL test used in this study included 90 items. It consisted of 3 sections; structure, written expression, and reading comprehension. The next instrument was a rating checklist including five questions about the components which are rated in analytical method of writing scoring. The checklist was used as treatment in the experimental group. That is, the students gave their instructors feedback on their strengths and weaknesses in writing during the study, and the instructor provided the necessary feedback to the students; therefore, they could overcome their weaknesses. The checklist assessed students' ability in five aspects of their writing ability including: introduction, body, and conclusion (organization); logical development of ideas (content); structure; mechanics; style and quality of expression. These aspects were supposed to be assessed by the students on a five point Likert scale. The teacher provided feedback at the bottom of the checklist. The third instrument was a writing test. The participants were instructed to write a 250 word composition on each of the following topics.

- a). The advantages and disadvantages of technology on man's life.
- b). Society should identify those children who have special talents and abilities and begin training them at an early age so that they can eventually excel in their areas of ability.

3.3 Procedure

In order to answer the main research questions, the following procedure was pursued. As a first step, the TOEFL test was administered to 105 language learners to make sure that control and experimental groups are homogeneous. Then, 73 language learners whose scores were between -1 SD below the mean and +1 SD above the mean were chosen as the subjects of the study. Next, the selected participants were randomly assigned to control and experimental groups.

After that, a descriptive topic was given to the sample to evaluate their writing proficiency. Having determined the experimental and control groups, the participants (in both groups) were asked to write a composition containing a maximum number of 150 words every three other sessions. The self-assessment techniques were utilized for the experimental group. The experimental group assessed themselves by the checklist. Each composition was corrected by three raters on the basis of the analytical method of scoring. Inter-rater and intra-rater reliability were calculated. Next, the average of the scores given by the three raters was determined. The teacher commented on the students' performance based on their weaknesses at the bottom of each checklist. Actually the comments were given to each participant individually and in the written form. However, the control group did not receive any feedback during the treatment. Finally, both groups were invited to write a composition as the post test. The means of the experimental group and the control group were compared through an independent sample t-test.

Data Analysis

To analyse the data, a set of statistical tests including Pearson Product coefficient, independent sample t-test, and KS test were run. The results of each part of the study are presented in the following parts of the study.

4. Results and Discussions

4.1 Results of the Pre-test

In order to confirm that the participants assigned to control and experimental groups were not initially different but homogeneous, an independent sample t-test was run. The results are shown in table 4.1.

As the results in the table 4.1 indicate, t value is 1.49 and sig.=.13 which is greater than .05. It means that the null hypothesis indicating that there is no significant difference between control and experimental groups is confirmed. Therefore, it was concluded that control and experimental groups were homogeneous at the outset of the study.

4.2 The Results of Reliability Estimates

In order to confirm the reliability of post-scores, both inter and intra-rater approaches of reliability estimates were applied. That is, the correlation coefficient between the three scores given by one particular rater was estimated; moreover, the scores given by three different raters were estimated. In tables 4.2 and 4.3, we present the correlation coefficient between the scores given by one rater at three different occasions as well as the scores given by three different raters.

As the results in the tables 4.2 and 4.3 indicate, the correlation coefficient between the three scores given by one rater and the three scores given by three different raters is significant at p value of .01 and correlation coefficient exceeds .95 which is a very high index for reliability.

4.3. Results of Post-tests

An independent t-test was run to compare the mean scores of the control and experimental groups on the post-test. As displayed in Table 4.4, the mean score for the control group is 13.88, with a standard deviation of 1.86. The mean score and the standard deviation for the experimental group are 17.40 and 1.41, respectively. The results show differences between the control and experimental groups' mean scores on the post-test.

The underlying assumption of independent t-test necessitates both groups enjoy homogenous variances. The Levene's test of equality of variance probes whether the two groups meet this assumption or not. The Levene's F value of 2.74 has a probability level of .103 (Table 4.5). Since the probability associated with the Levene statistic is greater than .05 level proposed by the researcher, it can be concluded that the two groups enjoy homogeneity of variances. Therefore, an independent sample t-test was run to compare the means of two groups. In table 4.5, the results of the t-test are shown.

The t-observed value calculated to compare the control and experimental groups' mean scores on the post-test is 8.5. This amount of t-value at 61 degrees of freedom is greater than the critical value of 2. Thus, it can be concluded that there is a significant difference between the control and experimental groups' mean scores on the post-test. Thus it can be concluded that the self-assessment treatment administered to the experimental group has a significant effect on the students' performance on the post-test of writing. The null-hypothesis as self-assessment does not improve the writing ability of the Iranian EFL learners is rejected.

5. Conclusion and Implications

The main purpose of this research was to investigate whether self-assessment influences Iranian EFL learners' writing skill. Oskarsson (1989) thinks that self-assessment can improve learning because it gives learners training in evaluation which is important to autonomous learning. To achieve the aim of this research, a TOEFL test (Appendix A) was administered to 105 advanced EFL students to ensure their homogeneity. Then 73 of them whose scores were between -1 SD below the mean and +1 SD above the mean were chosen as the subjects of the study. These students formed the experimental and control groups. A topic was assigned to determine the writing ability of the subjects in both groups. Having administered the pre-test, the experimental group was exposed to a variety of self-assessment techniques. Having finished the treatment, both groups attempted the same writing topics. As the results of data analysis indicated the null hypothesis of the study as self-assessment does not significantly impact Iranians EFL learners' writing skill was rejected. That is to say, administering self-assessment techniques to the experimental group improve writing skill significantly.

Besides that, it could be argued that self-assessment as a means of alternative assessment helps students to become autonomous learners and provide techniques for their own learning. The result of this study showed significant change in the experimental group; therefore, this study can be helpful for those teachers who like to help students develop strategies for their own learning, to help them find their strengths and weaknesses, and to

help them become more autonomous. Using self-assessment checklists also gives the teacher complete information about the students' progress and their failure in the process of learning.

In fact, the result of this study can be used in all educational centers. It has direct and indirect implications and applications in teaching, learning, and material development, syllabus design, and test development. Language teachers can benefit from self-assessment techniques in order to educate more active autonomous students who are at the same time better communicators.

Involving students in the assessment and evaluation process is an essential part of the balanced assessment. When students become partners in the learning process, they gain a better sense of themselves as readers, writers, and assessors. As students reflect on what they learn and on how they learn, they develop the tools to become learners that are more effective (Delmonte, 1997).

Self-assessment is rightly seen, as one of the pillars of learner autonomy. One of the fundamental elements of self-directed language learning is the opportunity given to learners to assess their own progress and thus help them to focus their own learning (Harris, 1997). It is widely accepted that self-assessment is a key learning strategy for autonomous language learning, enabling students to monitor their progress and relate learning to individual needs (Harris, 1997).

"It is important to help students become aware of the value of independent learning outside the classroom, so that they acquire the habit of learning continuously, and maintain it after they have completed their formal studies" (Lee, 1998, p.282).

Utilizing self-assessment techniques can improve writing ability as the result of this research showed. And self-assessment checklist helps the teacher to become familiar with the students' capabilities, deficiencies, and needs. Since the positive effect of self-assessment on writing skill was found in this study, methods of teaching and material development in language classes should be designed in such a way that encourage self-assessment techniques and consequently affect learners' writing skill.

Above all, the result of this study showed that there was some change in the subjects' level of writing ability in the experimental group and this change was significant. Thus, it can be concluded that self-assessment techniques help the students improve their writing ability. Finally, the researcher hopes the present study will be useful to those involved in the domain of language teaching to help students develop techniques for their own learning.

References

- Blanche, P. (1988). Self-assessment of foreign language skills: Implications for teachers and researchers. *RELC Journal* 19, 75-96.
- Blue, G. (1994). Self-assessment of foreign language skills: Does it work? *CLE working paper*, 3,18-35.
- Chen, Y. M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research*, 12(2), 235-262.
- Delmonte, R. (1997). Learning language with a "S. L. I. M." Automatic tutor. *Asiatica Venetiana* 2, 31-52. Retrieved December, 25, 2006 from: [http:// project. Cmg. Unive. It/slimpage/ slimasicc.pdf](http://project.Cmg.Unive.It/slimpage/slimasicc.pdf)
- Dickinson, L. (1987). *Self-instruction in language learning*. Cambridge: Cambridge University Press.
- Dikel, M. R. (2005). A guide to going online for self-assessment tools. Retrieved November 15, 2005, from: [http:// www.rileyguide.com/assess.html](http://www.rileyguide.com/assess.html)
- Erwin, T. D. (1991). *Assessing student learning and development: A guide to the principles, goals, and methods of determining college outcomes*. Hoboken, NJ: Jossey-Bass.
- Hamayan, E. V. (1995). Approaches to alternative assessment. *Annual Review of Applied Linguistics*, 15, 212-226.
- Harris, M. (1997). Self-assessment on language learning in formal setting. *ELT Journal*, 51(1), 12-14.
- Heilenman, L. K. (1990). Self-assessment of second language ability: The role of response effects. *Language Testing*, 7 (2), 174-201.
- Huerta-Macias, A. (1995). Alternative assessment: Responses to commonly asked questions. *TESOL Journal*, 5, 8-10.
- Hunt, J., Gow, L., & Barnes, P. (1989). Learner self-education and assessment: A tool for autonomy in the language learning classroom. *Language teaching and learning styles within and across cultures*, 207-217.

- Lee, I. (1998). Supporting greater autonomy in language learning. *ELT Journal*, 54 (4), 282-288.
- Mrudula, P. (2002). The influence of peer feedback on self and peer assessment of overall skills. *Language Testing*, 19 (2), 109-131.
- Mousavi, A. (1999). *A dictionary of language testing* (2nd ed.). Tehran: Rahnama Publication.
- Oskarsson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing*, 6(1), 1-13.
- Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge: Cambridge University Press.
- Richards, J. C., & Schmidt, R. (2002). *Dictionary of language teaching and applied linguistics* (3rd ed.). London: Pearson Education Limited.
- Rubin, J. (1975). What the “good language learner” can teach us. *TESOL Quarterly*, 9 (1), 41-51.
- Srimavin, W. & Daraswang, P. (2003). Developing self-assessment through journal writing. Retrieved November 15, 2005, from <http://www.independentlearning.org/ila03/ila03-srimavin-and-pornapit.pdf>

Table 4.1. T- test for Comparing the Mean Scores of the two Groups on the Pre-test

Group	N	Mean	SD	Levene's Test for Equality of Variances		t	Df	Sig.
Experimental	36	58.4	21.34	0.540	0.464	1.49	61	0.13
Control	27	51.9	23.06					

Table 4.2. Intra- rater Reliability Estimates

		First rating	Second rating	Third rating
First rating	Pearson Correlation	1	.988 **	.984 **
	Sig. (2-tailed)		.000	.000
	N	66	66	66
second rating	Pearson Correlation	.988 **	1	.983 **
	Sig. (2-tailed)	.000		.000
	N	66	66	66
third rating	Pearson Correlation	.984 **	.983 **	1
	Sig. (2-tailed)	.000	.000	
	N	66	66	66

** . Correlation is significant at the 0.01 level (2-tailed).

Table 4.3. Inter-rater Reliability Estimate

		Correlations		
		rater A	rater B	rater C
rater A	Pearson Correlation	1	.988**	.984**
	Sig. (2-tailed)		.000	.000
	N	66	66	66
rater B	Pearson Correlation	.988**	1	.983**
	Sig. (2-tailed)	.000		.000
	N	66	66	66
rater C	Pearson Correlation	.984**	.983**	1
	Sig. (2-tailed)	.000	.000	
	N	66	66	66

** . Correlation is significant at the 0.01 level (2-tailed).

Table 4.4. Descriptive Statistics Post-test for Control and Experimental Groups

	Group	N	Mean	Std. Deviation
Post test	Control	27	13.8827	1.86787
	Experimental	36	17.4074	1.41446

Table 4.5. Independent T-test Control and Experimental Groups on Post-test

	Levene's Test for Equality of Variances		t-test for Equality of Means			
	F	Sig.	t	df	Sig.	Mean Difference
Post test	2.746	.103	8.5	61	.000	3.52469