



The Effect of Self-Assessment on Iranian EFL Learners' Speaking Skill

Mahsa Ariafar (corresponding author)

ELT Department, Islamic Azad University, South Tehran Branch, Tehran, Iran

E-mail: Mahsa.ariafar@gmail.com

Hamid Reza Fatemipour

ELT Department, Islamic Azad University, Roudehen Branch, Tehran, Iran

Received: 26-02-2013

Accepted: 15-04-2013

Published: 01-07-2013

doi:10.7575/aiac.ijalel.v.2n.4p.7

URL: <http://dx.doi.org/10.7575/aiac.ijalel.v.2n.4p.7>

Abstract

Self-assessment is one of the attractive new alternative assessment forms which raises learners' awareness of their individual progress, and encourages learners to become autonomous learners (Harris, 1997). The present study aimed to investigate the effect of self-assessment on speaking skill of Iranian EFL learners. The participants were 60 pre intermediate Iranian EFL learners. The participants in experimental group received training and practiced self-assessment of their speaking ability during the semester; however, participants in control group experienced no self-assessment. At the end of the semester, the participants in both groups were interviewed as posttest. The results revealed that self-assessment practices helped the participants of the study to improve their speaking ability. Moreover, participants had generally positive attitudes toward the notion of self-assessment and showed their willingness to continue self-assessment in future; however, they were in favor of teacher assessment too. The findings of the study offer several implications for teachers and practitioners.

Keywords: Self-assessment, Speaking skill, EFL learners, Attitude

1. Introduction

In a learning process learners need to know what their abilities are, what their weaknesses are, how much progress they have made, and what they are able to do with the skills they have acquired. The current assessment practices in the Iranian TEFL context do not answer this need of learners. As Farhady (2003) states, "It has become a tradition in the Iranian educational system in general and language related majors in particular to implement traditional testing systems for the evaluation of the students' achievement" (p.1). The traditional assessment methods are very limited in providing opportunities for learners to reflect on their learning and to monitor their own progress critically. with the lack of chance for learners to self-reflect and without learners' awareness of their abilities, weaknesses, strengths, and the progress they have made, it would not be easy for them to learn efficiently and it may hinder the development of the desired skills and capabilities.

Harris (1997) suggests a solution for this situation and that is training learners in the skill of self-assessment. Self-assessment, which can be understood as measurement carried out by learners themselves, can play an important role in helping learners to monitor their learning progress or to judge their language proficiency or language ability. By involving learners in the assessment process, they learn the qualities needed for good performance, how to judge their own performance or achievement, how to set personal goals, and finally they develop the habit of self-reflection (Rolheiser and Ross, 1998). Also, if learners can do the self-assessment accurately enough, not only they don't have to depend completely on the teachers opinions, but also, they can make teachers aware of their individual needs.

Without a doubt, the notion of self-assessment is not a new concept in the assessment of language ability (Brindley & Scofield, 1989); however, how the EFL learners can be involved in the assessment of speaking skill and the effect of self-assessment on speaking skill of Iranian EFL learners still needs further investigation because a few studies have attempted to explore this area at the time of writing. Also, considering the fact that EFL context of self-assessment may be quite unfamiliar and challenging to many EFL learners (Harris, 1997), the researcher was interested to investigate the Iranian EFL learners' reaction to and attitude toward the self-assessment practices in order to better inform the self-assessment practices in the EFL contexts.

2. Background of the study

2.1 Involving learners in assessment

According to Saito(2003), "the current trends in learner-centered language teaching approaches, and a growing interest in instructiveness and authenticity have led to a greater interest in expanding the use of second language self-assessment" (p.1). It was claimed that self-assessment and peer assessment can help learners get to know 'how to learn' (Nunan, 1988). As Benson (2001) states, while assessing their language proficiency and progress, learners become

aware of their ability, strengths, and weaknesses. After recognizing their weaknesses, learners will then be able to seek help from their teachers in the areas in where they need support and improvement. Underhill (1987) introduces self-assessment as a test type which can be used for involving learners in the assessment of their spoken language, and refers to it as "the easiest, cheapest, and quickest form of assessment" (p. 22). He further argues that the learner is in the best position of assessing his/her speaking skill, and all learners have the ability to determine their oral proficiency; however, they lack the needed experience.

2.2 Self-assessment and learning

Rolheiser (1996) provides a theoretical model which justifies the way self-evaluation contributes to learning (as cited in Rolheiser and Ross, 1998). Based on this model, when learners evaluate their performance positively, self-evaluations encourage learners to set higher goals which result in committing more personal efforts. The combination of goals set by learners and the personal effort equals achievement. The achievement results in self-judgment which consequently results in self-reaction or responding to a judgmental question like "How I feel about that?" (Rolheiser & Ross, 1998, p. 4). Finally, all the goals, efforts, achievements, self-judgments, and self-reactions combine and affect self-confidence of the learner in a positive way. As Rolheiser and Ross (1998) argue, "Self-evaluation is really the combination of the self-judgment and self-reaction components of the model, and if we teach students to do this better we can contribute to an upward cycle of better learning" (p.4).



Figure 1. How self-assessment contributes to learning (adopted from Rolheiser and Ross, 1998)

2.3 Rationale for self-assessment

Blue (1994) believes that interest in self-assessment is the result of the general interest in the area of learner independence and autonomous learning. Self-assessment has been seen as "one of the more problematic areas of self-directed learning" and has always been criticized for reasons such as learners' lack of experience to make judgments in self-assessment (Coombe and Canning, 2002). Despite all the criticisms, there are some reasons why self-assessment practices should be encouraged in the EFL classrooms, especially in contexts where traditional assessment procedures are still dominant. Rolheiser and Ross (1998) argue that self-assessment enhances students' achievement and learning because of four reasons: "(i) self-evaluation will focus student attention on the objectives measured, (ii) the assessment provides teachers with information they would otherwise lack, (iii) students will pay more attention to the assessment, and (iv) student motivation will be enhanced" (p.10). They also believe that self-assessment is "unique" in asking learners to reflect on their performance (p. 10). Oscarson (1989) also mentions six reasons why self-assessment can be beneficial to language learning, including: promotion of learning, raised level of awareness, improved goal orientation, expansion of range of assessment, shared assessment burden, and beneficial postcourse effects.

2.4 Self-assessment techniques

Learners' ability to self-assess effectively develops over time and with experience (Cassidy, 2007). So, variety of tools and strategies may be used for the purpose of self-assessment considering various needs of learners. Some of the self-assessment strategies are: self-rating, portfolios, pre- and post-course writing, learner contracts, student progress card, learner diary, self-assessment questionnaire, videotapes, teacher-learner interviews, journals, and rating scale. These techniques need learners' awareness of their own progress, not only in terms of language but also of communicative objectives. According to Chamot and O'Malley (1994), "self-assessment requires the student to exercise a variety of learning strategies and higher order thinking skills that not only provide feedback to the student but also provide direction for future learning" (p.119).

2.5 Concerns about practicing self-assessment

A number of concerns have been raised regarding practicing self-assessment in the EFL contexts with EFL learners. Because of changes in teacher-learner relationships, the concept of self-assessment may be quite threatening and unfamiliar to many EFL learners (Blue, 1994). In this respect, learners support and training in self-assessment becomes necessary. Also, there is a controversy regarding learners' ability to self-assess accurately. Some studies suggest that learners have a tendency to say what they cannot do or the things that is difficult to do instead of what they are able to do (Bachman and Palmer, 1989). Moreover, it's likely that learners be able to assess their abilities more accurately when provided with specific self-assessment statements which are closely related to their personal experience (Ross,

1998). All of the above mentioned concerns have made the practice of self-assessment a challenging job in the EFL classrooms.

The study aimed to investigate the effect of self-assessment on speaking skill of Iranian EFL learners. The researcher was also interested to investigate the Iranian EFL learners' reaction to and attitude toward the self-assessment practices. To be able to meet the purposes of the study, the following research questions were posed:

Q1. Does self-assessment have any significant effect on Iranian EFL learners' speaking skill?

Q2. What is Iranian EFL learners' reaction to the self-assessment practices?

3. Method

3.1 Participants

The sample of the study consisted of 60 pre intermediate Iranian EFL learners who were selected out of 80 learners based on their scores on KET (Key English Test). All of the participants were male, native speakers of Persian language, and ranged in age mostly from 20 to 29 years old. The participants were chosen from available classes based on the convenience sampling technique, and were assigned to control (N=30) and experimental (N=30) groups. The participants in the experimental and control groups were all exposed to the same content and instructional method, and they had the same instructor.

3.2 Instruments

Multiple instruments were utilized for treatment and data collection purposes. The KET (Key English Test) was used as pretest for the purpose of testing the homogeneity of the participants in terms of their general English language proficiency. The KET was piloted among 30 EFL students; on the basis of their scores, an NRT item analysis (item facility and item discrimination) was performed. Since the results were according to the required range of item facility and item discrimination, no items were discarded. The reliability of the test was also calculated through Cronbach's Alpha and a value of .91 was obtained, indicating that the test was reliable and could be administered to the main participants of the study.

The other instrument of the study was Weir's analytic rating scale for assessing speaking skill which was used both for the treatment and data collection purposes. The rating scale assesses learners' ability in six aspects of the speaking ability, and consists of 6 criteria: 1) appropriateness, 2) adequacy of vocabulary for purpose, 3) grammatical accuracy, 4) intelligibility, 5) fluency, and 6) relevance and adequacy of content. Moreover, it should be mentioned that, at the end of the self-assessment form which contained the rating scale participants faced two questions which were related to their goal setting after the self-assessment practice. The questions were designed by the researcher to develop in learners more critical thinking of their performance, and were supposed to help learners make a plan for improving their skills or performance. In this way, participants were asked to think about their weaknesses, and if they were, or weren't, happy with their performance. The two questions are as follows:

- Am I happy with my performance?
- What do I need to improve on?

Oral interview test was the main research instrument in this study. Participants in both experimental and control groups were interviewed by two raters as both pre and posttests. For each test - the normal class teacher and another visiting examiner (who was also an instructor at the center and known to the students by sight if not by acquaintance) scored the speaking skill of participants separately using Weir's analytic rating scale. The inter rater reliability index for two raters who rated the speaking skill of participants on pretest and posttest were .84 ($P = .000 < .05$) and .85 ($P = .000 < .05$).

To elicit the speech production on the part of the participants, picture stories, or it's better to say comic strips (e.g. Tan Tan stories), were applied. The picture stories were described and narrated by the participants in both experimental and control groups. Picture stories were carefully selected so as to contain clear, easily recognizable features, and to be in accordance with the EFL learners' level of English language proficiency.

A questionnaire was also conducted and administered to elicit the participants' opinion and reaction to the self-assessment practices that they did during the treatment. The first draft of the questionnaire was examined by two EFL specialists and confirmed after some revisions. The questionnaire contained five questions of multiple choice type. The questions focused on information regarding participants' reaction, preferences, whether they were able to do self-assessment or not, the degree of learners difficulty doing self-assessment practices, willingness to continue self-assessment practices in future, whether they prefer teacher assessment or self-assessment, and learners' view regarding the subskill that was easy for them to self-assess.

3.3 Procedure

To conduct the research and to respond to the raised research questions, some stages were followed. At the beginning of the semester, to ensure the homogeneity of the participants, KET was administered to all of the participants who were 80 EFL learners. To decide on the final homogenous group based on the general English language proficiency, those learners whose scores were between one standard deviation below the mean and one standard deviation above the mean were chosen as the sample of the study. As the result, from among 80 English as foreign language learners, 60 of them were selected.

After testing the homogeneity of participants and assigning them to the control and experimental groups, a pretest was conducted in the form of oral interview test to participants in both experimental and control groups. Totally, the oral interview test lasted about 5 to 8 minutes, two raters were involved, and the atmosphere of the interview was friendly and unstressful.

Having examined the participants' speaking skill at the beginning of the study, the researcher followed various procedures in the experimental and control groups. The procedure followed for experimental group lasted about ten weeks and included four stages of self-assessment training, picture story narration, self-assessment practices, and answering attitude questions. Regarding self-assessment of productive skills such as speaking skill, Harris (1997) states that "it is essential to establish clear criteria for students to use when they assess their own performance" (p. 5). In the present study, since experimental group participants were neither knowledgeable enough to assess themselves, nor aware of the experts criteria for assessing speaking, they had a training session before the self-assessment practices. The purpose of training session was to facilitate learners' understanding of the assessment criteria. In the training session, learners assessed the recorded oral performances against the agreed criteria. The teacher then commented on the recorded sound to demonstrate her evaluation and scoring of the performances, and to explain how well the participants' evaluations had been made.

After the training session, each learner was given some pages of picture story to look at. Then, the teacher asked each learner to narrate the story and to speak freely about whatever the story narrates. The participants were allowed to add their own interpretation about the people or events involved, and were asked to tell the story or the sequence of events in one long stretch without any feedback from listeners or other participants in the same class. Moreover, the participants were asked to perform in a way that the listener can picture what being described, much as it would be in real life. The process of story narration happened in both experimental and control groups, and the just difference between the two groups was that participants in experimental group received training and self-assessed their speaking ability based on their narration of picture story. After narration of picture story, the participants in experimental group were provided with rating scales to self-assess their speaking ability. Soon after the last self-assessment practice, the participants in experimental group received the questionnaire asking them to report on their reaction to and perceptions of the self-assessment practices they did during the semester.

It is clear that for the control group no rating scale for self-assessment and no attitude questions were used; but oral interview test as pretest and posttest and picture stories were used for the task of narration. Finally, at the end of the course, all of the 60 participants in both groups of experimental and control took the oral interview test as posttest.

4. Results

In order to find answers to the research questions, some statistical analyses were performed. First, the assumption of normality was measured through the ratios of skewness and kurtosis over their respective standard errors. All the skewness and kurtosis ratios fell within the normality range of -1.96 and +1.96.

Then, an independent t-test was run to compare the experimental and control groups mean scores on the pretest of speaking skill in order to prove that the two groups enjoyed the same level of speaking ability prior to the main treatment. The descriptive statistics and inferential statistics have been displayed in tables 1 and 2. The result of the independent t-test ($t(58) = 1.371, P = .176 > .05$) showed that the amount of t-value at 58 degrees of freedom was lower than the critical value of 2. Thus, it was claimed that there was no significant difference between the treatment and control groups mean scores, and the two groups enjoyed the same level of speaking ability prior to the main study.

Table 2. Descriptive statistics on the pre test

	groups	N	Mean	Std. Deviation	Std. Error Mean
pretest	control group	30	10.5500	1.25499	.22913
	Experimental group	30	11.0000	1.28654	.23489

Table 3. T-test of the pre test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	.000	1.000	-1.371	58	.176	-.45000	.32813	-1.10683	.20683
Equal variances not assumed			-1.371	57.964	.176	-.45000	.32813	-1.10684	.20684

An independent T-test was also run to compare the mean scores of the treatment and control groups scores on posttest of speaking skill in order to probe the effect of self-assessment practices on participants speaking skill. The result of the independent t-test ($t(58) = 6.33, P = .000 < .05$) revealed that the amount of t-value at 58 degrees of freedom was higher than the critical value of 2. Thus, it was claimed that there was a significant difference between the treatment and control groups mean scores. Thus, the first null hypothesis which was: "Self-assessment does not have any significant effect on Iranian EFL learners' speaking skill" was rejected. That is to say, self-assessment practices helped the participants of the present study to improve their speaking skill as measured by oral interview test as posttest.

Table 4. Descriptive statistics on the post test

	group	N	Mean	Std. Deviation	Std. Error Mean
posttest	control group	30	10.6833	1.02118	.18644
	experimental group	30	12.2333	.86834	.15854

Table 5. T-test of the post test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	.685	.411	-6.333	58	.000	-1.55000	.24473	-2.03989	-1.06011
Equal variances not assumed			-6.333	56.539	.000	-1.55000	.24473	-2.04016	-1.05984

The second hypothesis assuming "it's not clear whether Iranian EFL learners respond positively or negatively to the self-assessment practices" was tested by estimating the frequencies and percentages of participants' answers to the questionnaire.

The first item of the questionnaire probes the participants' opinion regarding the helpfulness of the self-assessment practices. It was interesting that more than half of the participants (56.7 %) agreed that self-assessment was 'somewhat helpful', and less than half of participants (23.3 %) agreed that self-assessment was 'very helpful' for them. But, only 20 % of them saw self-assessment as 'not helpful'.

The second item of the questionnaire focuses on learners' preferences, whether they prefer self-assessment, teacher assessment, or they have no opinion. Half of the respondents (50 %) appeared in favor of teacher assessment. However, 30 % of respondents appeared to favor self-assessment, and 20 % of respondents had no opinion regarding both teacher assessment and self-assessment.

The third item of the questionnaire probes the degree of learners' difficulty doing self-assessment, and whether they were able of doing self-assessment or not. Regarding participants answers to this item, only a slender majority of respondents (6.7 %) regarded self-assessment as 'very easy' practice to do; however, half of the respondents (50%) considered it as 'easy'. Also, 36.7 % of respondents considered self-assessment as 'somewhat hard', and only 6.7 % of respondents considered self-assessment as very hard to do.

The fourth item of the questionnaire probes the participants' willingness to continue self-assessment practice in future. In answer to this item, half of the respondents (50 %) showed their willingness to continue self-assessment in future and answered 'yes'; however, only a slender majority of respondents (6.7 %) rejected the practice of self-assessment in future and chose 'No' as answer. Nevertheless, 30% of respondents were not sure about practicing self-assessment in future and answered 'maybe'; besides, just small percent of respondents (13.3 %) had no idea in this regard.

The last item of the questionnaire focuses on the subskill which was easier for participants to self-assess. The result of the analysis showed that the majority of respondents (33.3 %) saw fluency as an easy subskill to self-assess. However, 'vocabulary' was at the second place and 30 % of respondents chose it as an easy subskill. Besides, 20% of respondents regarded 'grammar' as an easy subskill, only 10 % of respondents chose 'intelligibility', a slender majority (3.3 %) of respondents selected 'appropriateness', and a similar majority (3.3 %) chose 'relevance and adequacy of content'.

5. Conclusion & Discussion

The main purpose of the study was to explore the possible improvement of Iranian EFL learners' speaking ability through self-assessment practices. Moreover, the researcher was interested to investigate the Iranian EFL learners' opinion and reaction to the self-assessment practices. With respect to the first research question of "Does self-assessment have any significant effect on Iranian EFL learners speaking skill?", a significant difference was observed between the means of the posttest scores in favor of the experimental group. The t-test analysis justified the rejection of the Null Hypothesis of "Self-assessment does not have any significant effect on Iranian EFL learners' speaking skill". Therefore, it was concluded that self-assessment practices during the semester helped the participants of the present study to improve their speaking ability.

The second research question addressed "What is Iranian EFL learners' reaction to the self-assessment practices?". The researcher was interested in learners' response to the self-assessment practices, whether negative or positive, to better inform the implementation of self-assessment practices in the TEFL context. After analyzing the answers to the questionnaire, it was concluded that the participants were generally positive towards the notion of self-assessment. The participants showed their willingness to continue self-assessment in future; however, they were in favor of teacher assessment. Moreover, they mostly considered self-assessment as an easy practice to do and accepted it as a somewhat helpful assessment.

With the support of the learners' favorable reactions and the positive effect of self-assessment on speaking skill of the participants, this study concludes that self-assessment is a viable alternative to involve students in the assessment process and serves as a way to diversify the assessment culture in Iran. The findings of the study may have some implications for various groups. Teachers can use the evaluative feedback resulted from self-assessment to increase their support of learners (as an example see Appendix). Teachers may also find it beneficial to adopt self-assessment as an effective means for development of foreign language learners' speaking skill. Also, by use of self-assessment practices, learners can identify the specific areas in where they need help and support, and then they can seek help from their teacher. Moreover, since the positive effect of self-assessment on EFL learners speaking skill was found in this study, methods of teaching and material development in language classes can be designed in such a way that encourage self-assessment techniques and consequently affect foreign language learners' language skill.

References

- Bachman, L., & Palmer, A. S. (1989). The construct validation of self-ratings of communicative language ability. *Language Testing*, 6, 14-25.
- Benson, P. (2001). *Teaching and Researching Autonomy in Language Learning*. London: Longman.
- Blue, G. M. (1994). Self-assessment of foreign language skills: Does it work? *CLE Working Papers*, 3, 18-35. Retrieved Aug, 23rd, from <http://www.eric.ed.gov>.
- Brindley, C., & Scoffield, S. (1998). Peer assessment in undergraduate programs. *Teaching in Higher Education*, 3(1), 79-89.
- Cassidy, S. (2007). Assessing "inexperienced" students' ability to self-assess: Exploring links with learning style and academic personal control. *Assessment and Evaluation in Higher Education*, 32(3), 313-330.
- Chamot, A.U. , & O' Malley, J. M. (1994). *The CALLA handbook: Implementing the Cognitive Academic Language Learning Approach*. White Plains, NY: Addison Wesley Longman.
- Coombe, C., & Canning C. (2002). Using self-assessment in the classroom: Rationale and suggested techniques. *Karen's Linguistics Issues* (February). Retrieved November 28, 2007, from <http://www3.telus.net/linguisticsissues/selfassess2.html>.
- Farhady, H. (2003). *Classroom Assessment: A Plea for Change*. Paper presented at the 2nd conference on issues in English language teaching in Iran, Tehran, Iran. Retrieved July 3rd, 2011, from <http://65.54.113.26/Publication/6146097/classroom-assessment-a-plea-for-change>
- Harris, M. (1997). Self-assessment of language learning in formal settings. *ELT Journal*, 51(1), 12-20.
- Nunan, D. (1988). *The Learner-centered Curriculum: A Study in Second Language Teaching*. Cambridge: Cambridge University Press.
- Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and implications. *RELC Journal*, 19(1), 75-93.
- Rolheiser, C., & Ross, J. A. (1998). Student self-evaluation: What research says and what practice shows. *Center for development and learning*. Retrieved December, 2011, from <http://www.cdl.org/>
- Ross, S. (1998). Self-assessment in language testing: A metaanalysis and analysis of experiential factors. *Language Testing*, 15(1), 120.
- Saito, Y. (2003). The use of self-assessment in second language assessment. *Working paper in TESOL and Applied Linguistics (Online)*. Retrieved August 2nd, 2008, from http://www.tc.columbia.edu/academic/tesol/WJFiles/pdf/Saito_Forum.pdf

Underhill, N. (1987). *Testing spoken language: A handbook of oral testing techniques*. Cambridge: Cambridge University Press.

Weir, C. (1993). *Understanding and Developing Language Tests*. New York: Prentice Hall.

Appendix

The assessment form completed by the students

Name: Shahriar

Use the rating scale for assessment of your speaking performance

Criteria	1. Not good 2. good 3. ok 4. very good
Grammar (Accuracy)	1 2 (3) 4
Vocabulary	1 (2) 3 4
Appropriateness	1 2 (3) 4
Intelligibility	1 (2) 3 4
Fluency	1 (2) 3 4
Relevance and Adequacy of Content	1 2 (3) 4

Total score:

Goal setting:

- Am I happy with my performance? no
- What do I need to improve on? Fluency -
vocabulary - Intelligibility