

Available online at www.sciencedirect.com

ScienceDirect

Procedia - Social and Behavioral Sciences 98 (2014) 1641 – 1649

Procedia
Social and Behavioral Sciences

International Conference on Current Trends in ELT

Self- assessment and Portfolio Production of Iranian EFL Learners

Roghayeh Sajedi*

Semnan University of Medical Sciences, Semnan, 3513138111, Iran

Abstract

The purpose of this study was to investigate the effect of self-assessment on Iranian EFL students' portfolio production. The experimental group was given the introduction to self-assessment which included a self-assessment guide, while the control group would not received the self-assessment guide and would not be exposed to any training during this research project. The experimental group outperformed the control group and the differences were statistically significant. Students engage in constant critical thinking and develop a strong sense of responsibility. For teachers, introducing self-assessment as a structured, formal practice will encourage students' involvement and lead to a highly learner-centered classroom.

© 2014 The Authors. Published by Elsevier Ltd. Open access under [CC BY-NC-ND license](https://creativecommons.org/licenses/by-nc-nd/4.0/).
Selection and peer-review under responsibility of Urmia University, Iran.

Keywords: self-assessment; portfolio; alternative assessments

1. Introduction

In recent years, there has been a notable increase in the amount of attention given to assessments in the classroom. Assessment of student learning has become a topic of great emphasis in the literature of education.

Many previously unfamiliar terms, such as authentic assessment, alternative assessment and portfolios, have become a part of the established vocabulary in educational publications. The main goals of the classroom assessment process appear to be changing, so that a stronger interest now lies in gathering data on students that focus on development over time, rather than comparing students to each other. An increasing amount of emphasis is also being placed on assessing the process by which students solve problem, as opposed to evaluating only the final outcome.

*Corresponding author. Tel.: +989122314631; fax: +982314453290.
E-mail address: sajedi4249@yahoo.com.

Today's focus seems to be on what students know and master, rather than on what they do not (Hargreaves, Earl & Schmidt, 2002). In fact, the shift in emphasis from teacher-centered toward student centered pedagogy has raised questions concerning the evaluation processes that would be best suited for this new approach.

Various novel assessment approaches, like portfolio assessment, peer assessment, and etc. have been proposed and are currently being utilized in the classroom (Taylor, 1994). It is inherent that, in a student centered setting, students become responsible for many of the decisions concerning what and how they want to learn. As it will be discussed in the coming sections, the learning of a language in a communicative approach requires that students be a part of the complete learning cycle, including assessment of language competency; here, it is believed that assessment "...should be done with learners, not to them" (Brindley, 2003, p. 316). The case in point is the use of alternative assessment techniques. Alternative forms in language assessment, e.g. self-assessment and peer assessment have been some of ways to involve students in the learning cycles. Alternative assessments can be regarded as testing methods that require students to create an answer or product that shows what they can do. These forms of assessment are in contrast to more traditional types of student assessment that focus mainly on the correct response (such as multiple-choice and short-answer type tests) (Brown & Glasner, 1999). Some of the evaluation methods mostly recognized in the field of alternative assessment include the following: self assessment, peer-assessment, portfolio assessment, protocol analysis, learning logs, journal entries, and dialogue journals.

1.1. Portfolio Assessment

Portfolio assessment is an approach to the assessment of students' writing that makes use of portfolios; a portfolio is "a collection of texts the writer has produced over a defined period of time to the specifications of a particular context" (Hamp-Lyons & Condon 2000, p. 261, cited in Hirvela & Sweetland, 2005, p.193). Portfolios have attracted considerable attentions among ESL writers and assessment specialists since the 1980s (Hirvela & Sweetland, 2005). Portfolios show how much students' writing has progressed from the beginning of the term.

Portfolio assessment, therefore, has the potential to create positive wash back on students' writing (Biggs & Tang, 2003; Hughes, 2003). Traditionally, students have been asked to write in a "one-draft, one-reader" context (Arndt, 1993). Having received a grade and minimal feedback from the teacher, students may make corrections on their drafts. After that, the learning process is supposedly finished and students are asked to write on another topic. The product approach to writing promotes students' reliance on a teacher's summative judgments rather than helping students to self-assess their own drafts before submission. The adoption of a portfolio approach in EFL writing classrooms may empower students' active participation in self-evaluating their own work within the writing process (Weigle, 2007).

Teachers respond to the materials not to provide an evaluation with a grade or score but to provide suggestions for revision as well as some general commentary about the individual's development as a writer. At the end of the course, teachers may also want to grade the portfolios which represent the results of what the students have learnt and the best work they have produced in the class (Mousavi, 1999, p.275). Yang (2003, p. 293) points out that portfolios raise "student's awareness of learning strategies," facilitate "their learning process", and "enhance their self-directed learning". *Conferencing* is an important component of portfolio assessment. It is a "semi-structured face-to-face conversation between a teacher and a student or a small group of students in which work being undertaken is discussed" (Richards & Schmidt, 2002, p. 105-106). Through conferencing, a teacher can respond to students' writing, and a student may discuss the problems he/she has encountered when he/she was writing.

In addition, conferencing provides teachers with the opportunity to give feedback on the problems discussed. Table 1 lists some of the benefits of portfolios.

Table 1: Benefits of Portfolios (adapted from Genesee & Upshur, 1996, p. 100)

<p><i>Portfolios provide:</i></p> <ul style="list-style-type: none"> •A continuous, cumulative record of language development. •A holistic view of student language learning. •Insights about progress of individual students. •Opportunities for collaborative assessment and goal-setting with students, educators, and other students. •Opportunities to use meta language to talk about language.
--

<p><i>Portfolios promote:</i></p> <ul style="list-style-type: none"> •Student involvement in assessment. •Responsibility for self-assessment. •Interaction with teachers, parents, and students about learning. •Student ownership of and responsibility for their own learning. •Excitement about learning. •Students' ability to think critically about schoolwork. •Collaborative, sharing classrooms.
--

1.1.1. What is Self-assessment?

Assessment for learning is the best counterpart for the emerging process of writing. Through portfolio and self-assessment learners are allowed to learn while checking the multifaceted process of writing skills.

Two recent movements in applied linguistics, including learner-centered language teaching and interest in authenticity, gave rise to the role of the learner or test-taker in the assessment process (Bachman, 2000). Self-assessment included in the process of selecting, reading and feedback can foster learners' reflection on their activities compiled in portfolio. Including self-assessment in the process of portfolio assessment will cover the lack of constructive feedback, which is observable in traditional achievement tests. "To encourage deep learning, teachers should give students an opportunity to engage in reflective dialogue and self-assessment". (Kathpalia and Heah, 2008). Self-assessment is considered as a particular type of metacognitive strategy, which aims to help students develop those characteristics of the 'good language learner' which involve the ability to assess their own performance and the ability to be 'self-critical' (Hedge, 2000). It is a procedure, which may involve the learner directly in the assessment process. Implicitly, it recognizes that learners should be able to take responsibility in making decisions about their own language learning development (Rea-Dickins, 2000). In fact, as Gardner (1996, p. 18) observes, "In self-access learning, students are encouraged to become more independent". The usefulness of self-assessment has been noted by various scholars (e.g., LeBlanc & Painchaud, 1985). The primary benefit of self-assessment is that it encourages students to become more actively involved in the educational process by requiring them to reflect on their own performances and by encouraging them to take greater responsibility for setting goals and making decisions about their own learning (Hughes, 2003; Dochy, Segers, & Sluusmans, 1999). Brown and Hudson (1998) mention that students' greater autonomy in self-assessment can substantially increase their motivation to learn the language. There are a number of ways in which self assessment is believed to improve learning processes. In her justification for self-assessment, Dickinson (1987) proposes three related reasons for the importance of self-assessment in learning. Firstly, Dickinson argues that assessment leading towards evaluation is an important educational objective in its own right, so training learners in this strategy is beneficial to learning. Secondly, self-assessment is a critical part of self-direction which will allow them to function effectively in the modern world; and, finally, in large classes, it is proposed that self assessment lessens the assessment burden on the

teachers by transferring greater responsibility to the students themselves. Self-assessment is desirable as it leads to more independent learning, and because it emphasizes the process of learning rather than the results, i.e. the product (Dickinson, 1987). Boud (1990) points out that self-assessment is a crucial element in an academic career, and have a meaningful role to play in other professions as well. Likewise, Cumming (1995) pointed out that self-assessment could encourage students to take greater charge of their writing skills. In this regard, students are more likely to get a wider perspective about different aspects of writing such as content, organization, mechanics, and rhetoric, when they self-evaluate their portfolio entries. Hamp-Lyons (as cited in Hirvela & Pierson, 2000) mentioned EFL students tend to underestimate their own writing. However, after participating in the portfolio program, students may realize that not only does self-assessment help them better diagnose their writing, but it also makes them respond to their work in a much more positive light.

2. Research questions

This study thus aims at investigating the effect of self-assessment on Iranian EFL students' portfolio production. This may be investigated via the following research question:

Do self-assessed portfolios have any effect on the writing skill of Iranian EFL learners?

3. Methodology

3.1. Participants

The study was conducted with 30 students taking an intensive course of English in Safir Institute. They were intended to take the Cambridge's FCE examination later after this course. To homogenize the participants, a mock exam (FCE) was administered. They were assigned into two classes. One of the classes was randomly selected as the control group and the other one as the experimental group.

3.2. Procedure

Having divided the participants into a control group and an experimental group based on their language proficiency scores, during a three session discussion basic elements of a good writing were introduced to both control and experimental groups. The purpose of doing so was to generate discussion and feedback from both groups and make sure that they all are aware of the fundamental requirements of a good writing. Following these sessions, the experimental group was given the Introduction to Self-assessment which included a Self-assessment Guide. To make sure that the control group would not receive the Self-assessment Guide and would not be exposed to any training during this research project, the sessions for meeting control and experimental groups were separated from each other. The raters were also provided with all the related materials.

During ten weeks, the participants wrote ten compositions from which the first and the last ones were scored as pre- and post-tests by three raters. The final score a participant received was the mean of the three raters' scores. Upon completion of a composition, students in the control group filed their writings in a cumulative fashion, while the students in the experimental group reflected and critically analyzed their own writing using the Self-assessment Guide. It has to be mentioned that the raters also used the same Guide for evaluating the students' compositions.

3.3. Materials

In this study, participants in the experimental group were asked to rate their own writing according to the guide. The guide uses a 4-point Likert scale that addresses different parts of the problem solving type of writing assignment. The four evaluation items used were: *strongly disagree*, *disagree*, *agree*, and *strongly agree*. In the Self-assessment Guide, neutral items like "not sure" or "no opinion" were excluded to encourage participants to make a real choice. Moreover, for each statement, extra space was provided for participants to write comments. A participants' final score is the sum of all the ratings for each item in the Self-assessment Guide.

3.4. Data analysis and results

As it was discussed in the previous sections, both experimental and control groups were supposed to write ten compositions from which the first and the last ones were scored as pre-and post-tests by the raters. In addition to the writing assignment, the participants in the experimental group had to score their own writing performances according to the criteria known as Self assessment Guide, which was later used by the three raters to rate the articles

of both groups. The Self assessment Guide was in fact a guide for marking the writing task. The items used in the Guide explore the degree of agreement/disagreement among the respondents concerning the statements used in the Guide. The Likert scale is used in the Guide where the responses are arranged on a scale of 1 to 4. The numbers in front of choices are the values they carry. The values are added up to find the respondents' overall scores. Maximum score that a respondent can get is 64. Having received the articles written by the both groups; three raters started rating the given articles according to the Self-assessment Guide. Descriptive Statistics of the experimental and control groups' writing performance manifests that control group's mean of rating is about 51.62 and mean of the experimental group's rating is 53 (Table 2). A comparison between experimental and control groups' makes it clear that there is not a big difference between these two means. Also, the difference between the standard deviations, that is the average of the differences of all scores from the mean or the average variability of all the scores around the mean (Riazi, 1999; Hatch & Farhady, 2002), does not seem to be significant and of course they are not so large (standard deviation of the experimental group is 4.24 and that of group is 4.11); that is, the distribution of both groups of the scores is close to the central point.

Table 2: Descriptive Statistics of the Raters' Ratings of the Experimental and Control Groups' Writing Performance in Pre-test

Descriptive Statistics					
Std. Deviation	Mean	Maximum	Minimum	N	
4.24439	53.0000	61.67	47.00	15	Raters' ratings of Exp. Group's performance
4.11299	51.6250	59.00	44.00	15	Raters' ratings of Cont. Group's performance

Having written the tenth composition, the raters started evaluating the writing as the post-test of the study. Descriptive statistics of the final scores obtained by the two groups indicates that the highest scores gained by the two groups does not differ from one another, but the lowest score gained by the control group is 16 points lower than that of the experimental group (Table 3; Figure 1). Moreover, the experimental group has a higher mean (53) and lower standard deviation (4.24) than the control group ($M = 48.08$; $S = 7.94$) (Table 3).

Table 3: Descriptive Statistics of the Raters' Ratings of the Experimental and Control Groups' Writing Performance in Post-test

Descriptive Statistics					
Std. Deviation	Mean	Maximum	Minimum	N	
8.55960	49.7500	62.00	41.00	15	1st Rater Cont.
8.11480	47.6250	61.00	39.00	15	2nd Rater
9.25833	46.8750	62.00	32.00	15	3rd Rater
				15	Valid N (list wise)
5.39096	52.5625	62.00	45.00	15	1st Rater Exp.
4.13471	53.1875	62.00	48.00	15	2nd Rater
4.05791	53.2500	62.00	47.00	15	3rd Rater

15	Valid N (list wise)
----	---------------------

Even though the mean score of the experimental group is higher than that of the control group, the mean difference by itself cannot reveal anything unless it is proved that the difference is not due to chance, or the difference is significant. In order to determine whether the difference between two groups' means is significant or not, independent samples *t*-test included in SPSS is used to check whether two groups are significantly different from one another. Here, in this section, the purpose of comparing the two means is to investigate the effect of self-assessment on the writing performance of the experimental group. Our hypotheses for testing the significance of the difference between the two groups' means will be as follows:

Our null hypothesis predicts that there is not a significant difference between the two groups' means. As the first step in comparing the two groups' means, descriptive statistics including the standard error of the difference between the two means is calculated (Table 4).

Table 4: Descriptive Statistics of the two Groups Final Ratings

Group Statistics			
<i>Std. Error Mean</i>	<i>Std. Deviation</i>	<i>Mean</i>	<i>N</i>
1.98525	7.94099	48.0833	15 <i>Control</i>
1.06110	4.24439	53.0000	15 <i>Experimental</i>

Using two samples *t*-test included in SPSS a comparison is made between the two groups' means. Also, the hypothesis of comparability of the variances is tested (Table 5).

Table5: *T*-test Results for Equality of the Experimental and Control Groups' Final Ratings

95% Confidence Interval of the Difference		T-test Equality of Means					Leven's Test for Equality of Variances		
Upper	Lower	Std. Error Difference	Mean Difference	Sig. (2-tailed)	df	t	Sig.	F	
-0.31945	9.51388	2.25103	-4.91667	0.037	28	-2.184	.014	6.820	Equal variances assumed
-0.25920	9.57413	2.25103	-4.91667	0.039	22.924	-2.184			Equal variances not assumed

Now we have the *t*-value (-2.184) and we have to compare it with the critical value and decide to reject the null hypothesis or retain it. To find the critical value, we consider the degrees of freedom and the significance level. The degrees of freedom are calculated by adding up the number of members in both groups and subtracting 2 from the sum ($df = (N1 - N2) - 2 = (15 + 15) - 2 = 28$). By referring to the table of *t*- distribution, it is found that with 28

degrees of freedom at the .05 level of significance, the critical value is 2.042. It can be easily concluded that because the t-value is greater than the critical value the difference between the two groups' means is significant. In other words, the null hypothesis is rejected (sig < .05). Rejection of the null hypothesis at the .05 level of significance also proves that there is a significant difference between the variances of the observations. To sum it up, rejection of the null hypothesis which predicted there is not a significant difference between the means of the two groups' final ratings proves the effect of self-assessment on the experimental group's writing performance. That is to say, self-assessment has had a positive influence in enhancing the writing ability of the experimental group.

4. Conclusion and pedagogical implications

This study sought to determine the effect of self-assessment on Iranian undergraduate students' portfolio production. According to the Self-assessment Guide, the experimental group outperformed the control group and the differences were statistically significant. Of course, the power of self-assessment could have been underrepresented in this study. Some of the participants stated that even though they read the Self-assessment Guide before writing, they did not revise their essays according to it because they wanted to identify their strengths and weaknesses from their original scores. However, hopefully, this study will raise students' and instructors' awareness of this powerful tool. Moreover, the study has demonstrated how students, teachers, as well as researchers may benefit from self-assessment. For students, in the process of self-assessment; they will achieve a clear understanding of the expected learning outcomes, find out their strengths and weaknesses, and set goals for future improvement. By assessing their own learning, students engage in constant critical thinking and develop a strong sense of responsibility. Learners are no longer passively receiving instructions from teachers but actively seeking ways to make further improvement. For teachers, introducing self-assessment as a structured, formal practice will encourage students' involvement and lead to a highly learner-centered classroom. Not only will teachers' teaching objectives and evaluation criteria be clearly presented and better understood, but also students' needs and goals will be attended to. Self-assessment under teachers' supervision will make teaching and learning integrated to fit the best interests of both parties. Finally, the strengths and weaknesses of this study will provide valuable references for future researchers.

References

- Arndt, V. (1993). Response to writing: Using feedback to inform the writing process. In M. Brock & L. Walters (Eds.), *Writing around the Pacific Rim* (pp.90-114). Clevedon, UK: Multilingual Matters.
- Bachman, L. F. (2000). Foreword. In G. Ekbatani and H. Pierson (Eds.), *Learner-Directed Assessment in ESL*, ix-xii. Mahwah, NJ: Lawrence Erlbaum Associates.
- Biggs, J. & Tang, C. (2003). Assessment by portfolio: Constructing learning and designing teaching. In P. Stimpson, P. Morris, Y. Fung, & Carr (Eds.), *Curriculum Learning and Assessment: The Hong Kong experience*. Hong Kong: Open University of Hong Kong Press.
- Brindley, G. (2003). Classroom-based assessment. In D. Nunan (Eds.), *Practical English Language Teaching* (pp.309-328). New York: McGraw-Hill.
- Brown, S. & Glasner, A. (1999). *Assessment Matters in Higher Education*. Buckingham: Open University Press.
- Brown, J. D. and Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-75.
- Boud, D.J. (1990). Assessment and the promotion of academic values. *Studies in Higher Education*, 15(1), 101-111.
- Cumming, A. (1995). Fostering writing expertise in ESL composition instruction: Modeling and evaluation. In: D. Belcher & G. Braine (Eds.), *Academic writing in a second language: Essays on research and pedagogy* (pp. 375-397). Norwood, NJ: Ablex.
- Dickinson, L. (1987). *Self-Instruction in Language Learning*. Cambridge: Cambridge University Press.
- Dochy, R, Segers, M., & Sluismans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24 (3), 331-350.
- Gardner, D. (1996). Self-assessment for self-access learners. *TESOL Journal*, 5 (3), 18-23.
- Hatch, E., & Farhady, H. (2002). *Research Design and Statistics for Applied Linguistics*. Tehran: Rahnama Publications.
- Hamp-Lyons, L. & Condon, W. (2000). *Assessing the portfolio: Principles for practice, theory and research*. Cresskill, NJ: Hampton Press.
- Hargreaves, A., Earl, L., & Schmidt, M. (2002). Perspectives on alternative assessment. *American Educational Research Journal*, 39(1), 69-95.
- Hedge, T. (2000). *Teaching and Learning in the Language Classroom*. Oxford: Oxford University Press.
- Hirvela, A. & Sweetland, Y. L. (2005). Two case studies of L2 writers' experiences a cross learning-directed portfolio contexts. *Assessing Writing*, 10, 192-213.
- Hughes, A. (2003). *Testing for language teachers*. (2nd ed.). Cambridge: Cambridge University Press.
- Hirvela, A. & Pierson, H. (2000). Portfolio: Vehicles for authentic self-assessment. In G. Ekbatani & H. Pierson (Eds.), *Learner-directed*

- assessment in ESL* (pp. 105-126). Manwah, N.J.: Lawrence Erlbaum Associates.
- LeBlanc, R., & Painchaud, G. (1985). Self-assessment as a second language placement instrument. *TESOL Quarterly*, 19, 673-687.
- Mousavi, S. A. (1999). *A Dictionary of Language Testing*. 2nded. Tehran:Rahnama.
- Kathpalia, S. & Heah, C. (2008). Reflective writing: Insights into what lies beneath. *Relc Journal*. 39(3), 300-317.
- Rea-Dickins, P. (2000). Classroom assessment. In T. Hedge (Ed.), *Teaching and Learning in the Language Classroom* (pp. 375-401). Oxford: Oxford University Press.
- Riazi, A. M. (1999). *A Dictionary of Research Methods: Quantitative and Qualitative*. Tehran: Rahnama.
- Taylor, C. (1994). Assessment for measurement or standards: The peril and promise of large-scale assessment reform. *American Educational Research Journal*, 31(2), 231-262.
- Weigle, S. C. (2007). Teaching writing teachers about assessment. *Journal of Second Language Writing*, 16(3), 194-209
- Yang, N. (2003). Integrating portfolios into learning strategy-based instruction for EFL college students. *IRAL*, 41, 293-317.
- Richards, J. C. & Schmidt, R. (2002). *Dictionary of Language Teaching and Applied Linguistics*. (3rd ed.). Longman.

Appendix

Self-assessment Guide for Problem-Solving Type of Writing Code:

Please underline ONE choice out of the four. You are welcome to provide additional comments.

Strongly Disagree = 1 Disagree = 2

Agree = 3 Strongly Agree = 4

A. Organization

1. Background information about the problem is first provided.

1 2 3 4

Comments:

2. The problem is then identified with details.

1 2 3 4

Comments:

3. Solution(s) is (are) then proposed to solve the problem.

1 2 3 4

Comments:

4. The author evaluates the solution(s) at the end.

1 2 3 4

Comments:

B. Content and Structure

5. A variety of vocabulary is used correctly.

1 2 3 4

Comments:

6. Various sentence structures have been used in the essay.

1 2 3 4

Comments:

7. Transition words or expressions have been used to connect ideas.

1 2 3 4

Comments:

8. Main ideas are expressed in key sentences.

1 2 3 4

Comments:

9. The supporting details to the problem are relevant.

1 2 3 4

Comments:

10. The supporting details to the problem are sufficient.

1 2 3 4

Comments:

11. The evaluation of the solution(s) is fair.

1 2 3 4

Comments:

C. Grammar

12. There are no spellings or typo mistakes in the essay.

1 2 3 4

Comments:

13. Punctuation has been used correctly in the essay.

1 2 3 4

Comments:

14. Articles have been used correctly in the essay.

1 2 3 4

Comments:

15. Passive voice has been used frequently and correctly.

1 2 3 4

Comments:

16. There is no obvious influence of a language other than English.

1 2 3 4

Comments:

D. Other

Total score: