

## The Effect of Self-Assessment on EFL Learners' Receptive Skills (Kesan Penilaian Kendiri terhadap Kemahiran Reseptif dalam Kalangan Pelajar Bahasa Inggeris sebagai Bahasa Kedua (ELF))

GOUDARZ ALIBAKHSHI\* & HASSAN SHARAKIPOUR

### ABSTRACT

*Past research has shown that self-assessment influences language learners' speaking and writing performance. However, few studies have examined the impact of self-assessment on EFL learners' listening and reading abilities. This study was an attempt to investigate any effect of self-assessment on beginner and intermediate level proficiency EFL learners' receptive skills. The subjects were 120 Iranian language learners selected through convenience sampling and divided into beginner and intermediate level proficiency. The experimental group was exposed to self-assessment and assessed their learning. Both the experimental and control groups received a posttest. Data were analyzed using a 2-way ANCOVA test. The results showed that self-assessment significantly improved EFL learners' receptive skills. However, the effect of self-assessment for listening skill was less than for reading skill. Also, the effect was greater for the intermediate group compared to the beginner group. An implication of the study is that self-assessment can possibly be used as an alternative to traditional teacher-made tests to help language learners to further develop their receptive skills.*

*Key words: Self-assessment, Language learners, EFL, Receptive skills*

### ABSTRAK

*Kajian lepas menunjukkan bahawa penilaian sendiri mempengaruhi kemahiran bertutur dan menulis pelajar. Namun, terdapat juga kajian lepas yang pernah mengkaji kesan penilaian sendiri dari segi kemahiran mendengar dan membaca dalam kalangan pelajar EFL. Kajian ini dijalankan untuk mengenal pasti kesan penilaian sendiri pada tahap permulaan dan tahap sederhana pelajar EFL. Responden kajian ini terdiri daripada 120 orang pelajar Iran yang dipilih berdasarkan pensampelan selesa yang dipecahkan kepada pelajar tahap permulaan dan tahap sederhana dan menilai tahap pembelajaran. Kedua-dua kumpulan eksperimen dan kumpulan kawalan diberi ujian pasca. Data dianalisis dengan menggunakan kajian ANOVA. Hasil kajian ini menunjukkan bahawa penilaian sendiri memang mempengaruhi pelajar dan kesannya adalah lebih baik berbanding dengan kumpulan pelajar tahap permulaan. Penilaian sendiri boleh digunakan sebagai cara pengajaran dalam bilik darjah untuk membantu pelajar menambah baik kemahiran reseptif.*

*Kata kunci: Penilaian diri, Pembelajaran bahasa, English sebagai bahasa asing (EFL), Kemahiran reseptif*

### INTRODUCTION

Assessment is an integral part of every educational system through which learners' achievements are evaluated on the basis of the material taught (Stefani 1998; Orsmond & Taras 2001; Patri 2002; Alderson 2005; Kato 2009; Buler & Lee 2010; Abbaszadeh 2012; Cheng & Curtis 2012). In learner-centered pedagogy, learners are active participants in education and learning and the task of evaluation or assessment is given to the students. Students take part in the process of evaluation of their learning termed "self-assessment". Bailey (1998) defines self-assessment as "procedures by which the learners themselves evaluate their language skills and knowledge."

The focus in the self-assessment approach is on the active participation of learners in the learning and evaluation process. This participation motivates learners for they see themselves as responsible for their own learning. It also provides learners with autonomy which in

the long term makes them life-long independent learners. Consequently, learners will be able to make judgments about their own learning and identify their own weaknesses and strengths. This further raises their awareness of their own state of knowledge.

Oscarson (1997, 1989) believes in the importance of learners' responsibility and argues that assessment is not the sole responsibility of the teacher, rather it is a mutual responsibility of both learners and teachers. This mutual responsibility will lead to a democratic educational system and learning context. Both learners and teachers and the institution will benefit from the practice of self-assessment. Various scholars (Freeman & Lewis 1998; Boud 2000; Patri 2002; Alderson 2005) believe that self-assessment helps learners in learning language, but it is unclear whether it has a significant impact on EFL learners' receptive skills. In other words, it is uncertain whether self-assessment has a positive or negative impact on learners' listening and reading skills. This research attempts to study the effect of

self-assessment on the listening and reading performance of beginner and intermediate EFL learners.

#### RESEARCH QUESTIONS AND HYPOTHESES

The main objective of the study was to determine whether self-assessment has any significant effect on beginner and intermediate EFL learners' listening and reading skill performance. To address the purpose of the study, the following research questions were posited:

Q1: Does self-assessment significantly affect EFL learners' reading performance?

Hypothesis: Self-assessment does not significantly affect EFL learners' reading performance.

Q2: Does self-assessment significantly affect EFL learners' listening performance?

Hypothesis: Self-assessment does not significantly affect EFL learners' listening performance.

#### REVIEW OF RELATED LITERATURE

A great number of research in the 1980s was concerned with the development of self-assessment instruments and their validation (Oscarson 1997; Lewkowicz & Moon 1985). As a result, many approaches including pupil progress cards, learning diaries, log books, rating scales and questionnaires were developed (Boud 1986; Dearing 1997; Falchikov 1997; Stefani 1998; Taras 2001, 2002). Self-assessment helps learners to become autonomous and to be aware of their learning and to reflect on their development (Freeman & Lewis 1998; Boud 2000). McDonald and Boud (2002, 2003) found that when learners assess their own learning, their learning will be highly promoted. Higgins et al. (2001) and Ivanic et al. (2000) stated that for the development of self-regulation in learners, both teacher and learner feedback on the learning process are needed and that this ability to give feedback must be promoted in both teachers and learners. Jewah et al. (2004) concurs that learners must occupy an important and active role in the process of giving feedback.

Rowntree (1987), Boud (2000) and Tares (2001) stated that the use of self-assessment in some areas such as in England and Wales is probably uncommon because it is seen as inconsistent with the conventional forms of assessments. Similarly, Carton (1993) discussed how self-assessment can become part of the learning process. He describes his use of questionnaires to encourage learners to reflect on their learning objectives and their preferred modes of learning. He also presents an approach to monitoring learning that involves the learners in devising their own criteria; an approach that he believes helps learners to become more aware of their own cognitive processes.

Likewise, Butler and Li (2010) investigated the effectiveness of self-assessment among young EFL learners.

They found some positive effects of self-assessment on the students' English language performance as well as on the students' confidence in learning English although the effect sizes were small. The study also found that teachers' and students' perceptions of the effectiveness of self-assessment are different depending on their teaching/learning contexts. A number of interesting insights were also discovered through interviews with teachers. The teachers were asked about the best way to utilize self-assessment as part of foreign language instruction in contexts whereby teacher-centered teaching has been traditionally valued.

Alderson (2005) investigated the importance of self-evaluation in the second language classroom and stressed the advantages of having students to keep a regular journal. Taking the methodological framework offered by the Communicative Approach to language teaching as a starting point, the dynamic inter-dependence of purpose, methodology and evaluation within the curriculum were studied. In this sense, formative or ongoing evaluation becomes one of the most practical assessment techniques for controlling students' progress as well as the effectiveness of a teaching program. Self-evaluation has affective advantageous in that students assess their own learning processes and participate in classroom management.

A number of studies have found that training and feedback influence the accuracy of self-assessment (Orsmond & Taras 2001; Patri 2002; Stefani 1998). Adequate training before doing self-assessment is believed to be effective (Hanrahan & Isaacs 2001; Li 2001; Taras 2002). Some researchers provide recommendations for effective self-assessments. Lejk and Wyvill (2001), for example, recommend a holistic approach rather than a category-based approach. Blatchford (1997) found that there is a significant association between self-assessments and attainments in both English reading and mathematics. Taras (2001) reported that the experience of active participation of learners and the teachers will enhance the process of self-assessment.

#### APPROACHES TO SELF-ASSESSMENT

Different approaches to self-assessment have been investigated. For example, Fallows and Balasubramanian (2001) reported that compulsory training combined with multiple ratings offer many benefits. Motivation also plays a significant role in the accuracy of self-assessment. AlFallay (2001) concludes that learners who have integrative motivation do more accurately in assessing themselves than those with instrumental motivation. The former group was also seen to be less apt to reflect overestimation than the latter group. Furthermore, he claims that language proficiency also influences the accuracy of self-assessment inducing those with higher proficiency to be more accurate than those with lower proficiency. High proficiency learners had to some extent, underestimated their performance, while the lower proficiency level learners often overestimated their performance.

In addition, researchers claim that motivation intensity is an important factor in the accuracy of self-assessment (Livesey 1992; Morton et al. 1999). A positive relationship between self-esteem and some other personality traits has also been reported (Calderon 1991; Collins 1993; Lindholm-Leary & Borsato 2002). In their study, Heilenman (1991) and Wesche et al. (1990) observed that learners are able to self-assess their achievements accurately. Stankowv (1998) reported that students are often over-confident on the tests of vocabulary and general knowledge.

Brantmeier (2005a, 2005b, 2006) reported that levels of self-assessed abilities positively correlated with levels of enjoyment. The study showed significant effects for self-assessed ability and enjoyment on written recall, but no effects were reported on multiple-choice questions. These studies lend support to the hypothesis that self-assessment can be accurate for placement. Oscarson (1997) claimed that there seems to be a fairly common consensus that the question of accuracy and appropriateness of self-estimates of proficiency depends to a great extent on feature of context and on the intended purpose of the assessment for each individual case.

Harutyunyan and Gasparyan (2003) investigated the possibility of integrating students' self-assessment into the evaluation process of the Intensive English Program (IEP) for students at the American University of Armenia (AUA). The purpose was to raise the students' awareness of their strengths and weaknesses in different language learning areas and to prepare the students for autonomous English language learning. Most studies on self-assessments have involved older subjects such as college students (Falchikov & Boud 1989; Topping 1998; Falchikov & Goldfinch 2000) and in-service staff (Saavedra & Kwun 1993; Jones & Fletcher 2002). However, little research has focused on the effects of self and peer-assessments in primary and middle schools. Significant differences between the characteristics of adolescents and adults suggest that studies should specifically investigate whether self- and peer-assessment are suitable for younger students.

#### RELIABILITY AND VALIDITY OF SELF-ASSESSMENT

According to Matsuno (2009), many researchers have reported high correlations in the assessments of students and teachers, while other studies reported low correlations between student and teacher assessments. Pierce et al. (1993) was based on school aged learners in a French immersion program in Canada. Learners assessed themselves against two criteria: by comparing themselves with a native speaker and by reflecting on the difficulty they experienced with everyday tasks in French. Results were compared against learners' results on proficiency tests of the four skills. The researchers concluded that self-assessment is not a reliable indicator of proficiency. They pointed out that many of the subjects had little or no access to the target language or to native speakers outside the classroom. Hence, it was probably difficult

for the learners to imagine how they would perform. In a comparison of a test of Dutch as a second language for adult learners and a self-assessed version of the same test, Janssen-van-Dieten (1989) found the self-assessed version less reliable although earlier studies and her pilot studies were more encouraging. For the researcher, the value of self-assessment is in "its positive influence on the learning process" (Janssen-van-Dieten 1989). Thomson (1996), in studying learners of Japanese as a foreign language, also felt positive about using self-assessment despite finding considerable diversity in the accuracy of self-marking.

Other studies have however reported self-assessments as reliable. Bachman and Palmer (1989) found that members of a multilingual, multicultural group of adult learners of English as a Foreign Language (EFL) in the US were able to reliably self-rate themselves for their communicative language abilities. Another example of success with self-assessment was reported in Blanche's (1990) study whereby a group of adult learners of French as a Foreign Language were able to estimate their own speaking ability. He concluded that "the overall accuracy of the self-evaluation... is impressive" (Blanche 1990). Variability in sample size, age of subjects, cultural and educational backgrounds, target language, the test format, the education context and the criteria against which self-assessment is compared could all affect reliability. What is comforting is that even when the results are against reliability, researchers maintain the value of self-assessment. One way to validate individual self-assessments is for the teacher to randomly check some of the results. This would encourage learners to be honest and realistic in their self-marking and would contribute to accreditation. In addition, regular random checking would provide a clearer understanding of the reliability issue (Gardner and Miller 1999). Xiao and Lucking (2008) examined the validity and reliability of student-generated assessment scores. The findings indicated that the validity and reliability of student-generated rating scores were high. Alfallay (2004) investigated the role of psychological and personality traits of EFL learners for accuracy in their self- and peer-assessments. The study revealed that long periods of practice and sufficient feedback had a positive effect on the accuracy of self-assessment. He asserts that students with low self-esteem are the most accurate in assessing their performance, whereas learners with instrumental motivation are the least accurate (Alfallay 2004). Sung et al. (2005) showed that significant consistency is found between the results of student self- and peer-assessments and that of teacher assessments.

Blaska and Krekeler (2008) investigated the reliability of self-assessments of pronunciation skills and attempted to understand the causes of difficulties. In the study, 46 advanced learners of German assessed their own articulation of different speech sounds in comparison with the sounds produced by a native speaker. In 85% of all cases, the assessments of the raters and the self-assessments were identical. However, the learners only identified half

of the number of speech sounds which the raters believed to be inaccurate. The conclusion of the study is that even experienced L2 learners seem to find it difficult to self-assess correctly their pronunciation skills.

Oscarson (1997) summed up progress in the area of self-assessment with the reminder that research in self assessments is fairly new. He concluded that there are still many problems that need to be addressed. For instance, learner goals and interpretations need to be adjusted with external necessities. In addition, self-assessment is not self-explanatory; it must be introduced slowly and learners need to be guided and supported in their use of the instruments.

#### SUBJECTS

The subjects of this study were 120 Iranian EFL learners majoring in English translation and Teaching English as a Foreign language (TEFL). The participants, with their ages ranging from 20-25, and without control for gender, were selected through convenience sampling from among the language learners in the Islamic Azad University at Rudehen, Tehran North, and Karaj branches based on the results of a placement test administered. Sixty of the participants were beginner in terms of language proficiency because their scores on the entrance examination to university were below 25%. In addition, 60 junior students who were taking reading and speaking courses were selected as intermediate language learners. The students from the two levels of proficiency were divided into two groups: control and experimental. Learning achievement of the experimental group was self-assessed by the learners themselves. However, the control groups' achievement were assessed by the teachers. At the end of the semester, all the participants took a post test which comprised reading and listening parts.

#### PLACEMENT TEST

The placement test is a kind of test used to measure the general language ability of the learners in different skills. The placement test used in this study was general language test consisting of 80 items adapted from Longman paper and pencil tests. It contained reading and listening parts. Reliability of the test was administered through KR-21. The reliability index of the test was .78 which seemed to be acceptable. The students were grouped according to proficiency levels based on this placement test.

#### POST-TEST

This test consisted of two parts: listening (40 items) and reading (40 items). This test was also adapted from the Longman paper and pencil TOEFL test. The participants

of each group sat for the test under the same conditions (time and rubrics). The reliability of the test was estimated through KR-21 approach and the reliability index was found to be at an acceptable level of 0.75.

#### PROCEDURE

The selected subjects were divided into control and experimental groups. For listening courses taken by both intermediate and beginner participants, the teachers were requested to make use of the teaching techniques in the course listening textbooks. For reading courses, the teachers used two different textbooks for extensive reading and one common textbook for intensive reading (Peterson reading flash book). In the reading course, the main focus was on teaching reading skills to the language learners. The only difference between the control and experimental groups was the way they were assessed. Experimental groups received treatment on self-assessment and they were given instructions to self-assess their listening and reading using self-assessment report sheets. During the treatment period, different techniques of self-assessment were introduced by the teacher. The teacher defined each technique in detail and asked learners to practice the technique for the next class sessions. Whenever necessary, the teacher provided the learners with the necessary information and consultation. At the beginning of the training, the teacher gave support to each step taken by the learners and as the learners become more and more proficient in using the self-assessment techniques, the teachers' support became less to make the learners more autonomous and independent. After a 15 to 90-minute treatment all groups received the same listening and reading test.

#### DATA ANALYSIS

Due to the design of the study, two pretests, two dependent variables, and two independent variables (listening and reading posttest), and two groups (control and experimental), the analysis of covariance was used because it permits researchers to statistically control for differences on the pretest so that posttest differences would not be due to initial differences before training. As listening and reading posttests were treated as two different dependent variables, we had to run two-way ANCOVAs, one for beginner groups and one for intermediate groups.

#### RESULTS OF THE STUDY

In the ANCOVA test, the scores on the pre-test served as the covariate to 'control' for pre-existing differences between the control and experimental groups. The assumptions of ANCOVA including the assumptions of unequal sample sizes, normal distribution of the scores,

homogeneity of variance, linearity, and homogeneity of regression and reliability of covariates were all checked and the researcher was sure that the assumptions were not violated. The following section provides information on

the key assumptions of the study including homogeneity of variances and homogeneity of regression as well as the descriptive statistics and inferential statistics.

TABLE 1. Homogeneity of regression

| Source                | Dependent Variable | Sum of Squares | Df | Mean Square | F   | Sig. | Partial Eta Squared |
|-----------------------|--------------------|----------------|----|-------------|-----|------|---------------------|
| Beginner groups       | Listening pretest  | .371           | 1  | .371        | .21 | .64  | .004                |
|                       | Reading pretest    | 5.387          | 1  | 5.387       | 3.  | .08  | .054                |
| Intermediate Groups * | Listening pretest  | 1.580          | 1  | 1.580       | .93 | .33  | .017                |
|                       | Reading pretest    | .147           | 1  | .147        | .08 | .77  | .002                |

In the output, the value of interest for this study is the significance level of the interaction (e.g. shown in the table above as Group\*listening pretest). The significance level for the interaction between beginner groups and both listening and reading pretest is greater than .05 (.64 and .08), hence, interaction is not statistically significant indicating a non-violation of the assumption. The results also showed that the interaction between intermediate groups and both listening and reading pretests is non-significant (significance value = .33 and .77, respectively). Therefore, it could be suggested that the difference between the control and experimental groups on the dependent variables is due to the treatment (self-assessment).

TABLE 2. Levene's test of equality of error variances

| Test                   | F     | df1 | df2 | Sig. |
|------------------------|-------|-----|-----|------|
| Beginner listening     | 2.285 | 1   | 58  | .136 |
| Beginner reading       | 5.475 | 1   | 58  | .09  |
| Intermediate listening | 1.285 | 1   | 58  | .15  |
| Intermediate reading   | 7.475 | 1   | 58  | .23  |

Levene's Test of Equality of Error Variances was used to check for homogeneity of variance between the control and experimental groups. As can be seen in Table 2, the significant values (p) of 0.136, .09, .15, and .23 are greater than 0.05; therefore, the null hypotheses were not rejected. There was no significant difference between the groups and the variances remained equal; thus the assumption of ANCOVA was not violated.

As shown in Table 3, the mean scores of beginner control and experimental groups on the listening posttest are 17.1 and 20.1 with standard deviations of 2.32 and 2.39, respectively. Therefore, the mean score of the beginner experimental group is higher than that of the control group. The table also indicates that mean scores of the control and experimental groups (beginners) on the reading posttest are 19.43 and 24.56, respectively.

TABLE 3. Descriptive statistics of intermediate and beginner groups for listening and reading

| Groups                    | Tests     | Mean  | SD   | N  |
|---------------------------|-----------|-------|------|----|
| Beginner Control          | Listening | 17.1  | 2.32 | 30 |
|                           | Reading   | 19.43 | 2.21 | 30 |
| Beginner Experimental     | Listening | 20.1  | 2.39 | 30 |
|                           | Reading   | 24.56 | 2.23 | 30 |
| Intermediate control      | Listening | 24.1  | 2.15 | 30 |
|                           | Reading   | 25.8  | 2.17 | 30 |
| Intermediate experiential | Listening | 27.3  | 2.14 | 30 |
|                           | Reading   | 33    | 2.47 | 30 |

In addition, the results in Table 3 indicate that the mean scores of intermediate experimental groups on listening and reading are 27.3 and 33, respectively. However, the mean scores of the control groups on listening and reading tests are shown to be at 24.1 and 25.8, respectively. A two-way ANOVA test was run to check whether the difference between the control and experimental groups was significant. The results are shown in Table 4.

TABLE 4. The results of ANCOVA for beginner and intermediate learners

| Source              | Dependent Variable | Type III Sum of Squares | Df | Mean Square | F     | Sig. | Partial Eta Squared |
|---------------------|--------------------|-------------------------|----|-------------|-------|------|---------------------|
| Beginner groups     | Reading            | 246.214                 | 1  | 246.214     | 147.9 | .000 | .70                 |
|                     | Listening          | 198.705                 | 1  | 198.705     | 109   | .000 | .60                 |
| Intermediate groups | Listening          | 195.797                 | 1  | 195.797     | 116.6 | .000 | .67                 |
|                     | Reading            | 773.096                 | 1  | 773.096     | 160.8 | .000 | .74                 |

As observed in Table 4, the difference between the mean scores of beginner language learners on both listening ( $F = 147.9$ , significant value = .000) and reading tests ( $F = 109$ , sig = .000) are significant. The results also indicate that the partial Eta squared of reading was shown to be .70, but that of listening was .60 which is smaller than the effect of the reading test. Thus, the null hypothesis of the study was rejected at p value of less than .001 and the alternative hypothesis was accepted. In other words, self assessment seems to have significantly improved beginner language learners' receptive skills.

Moreover, the results in Table 4 (second column) show that the difference between intermediate control and experimental groups on listening test was significant ( $F = 116.6$ , sig. = .000, and Eta = .67). Also, the difference between control and experimental groups on reading was significant ( $F = 160.8$ , sig. = .000, Eta = .74). Therefore, the null hypotheses which state that self assessment does not significantly improve intermediate language learners' reading and listening were safely rejected.

#### DISCUSSION

The present study tested the hypotheses that self assessment does not significantly improve Iranian beginner and intermediate EFL learners' listening and speaking skills. In doing so, participants in experimental groups in both beginner and intermediate level received a 15 session treatment in which they got familiar with self-assessment and its techniques. In addition, they learned how to apply self-assessment in the process of their language learning particularly in reading and listening skills.

The data of the study for both beginner and control participants were analyzed through ACNOVA. From the two-way ACNOVA, several interesting findings were revealed. First, the results of the study showed that there was a significant difference between the mean scores of the beginner participants on both listening and reading posttests. The mean of the experimental groups were significantly higher than the mean scores of the control groups. Therefore, it could be strongly argued that self assessment had significant impact on beginner language learners' reading and listening performance. The findings are therefore consistent with the findings of Stefani 1998; Orsmond et al. 2000; Taras 2001; Patri 2002; Hanrahan & Isaacs 2001; Li 2001; Smith et al. 2002; Taras 2002; ALfallay 2004; Sung et al. 2005; Dlaska & Krekeler 2008).

The results are also consistent with the findings of Abbasszadeh (2012) who found that self assessment significantly improves speaking and writing performance. However, they believed that self assessment has the same impact on both intermediate and beginner language learners.

The results also revealed that the Partial Eta Squared of the reading posttest for beginner participants was .70;

whereas, the Partial Eta squared of the listening test was .60. This value also indicates how much of the variance in the dependent variable is explained by the independent variable. When partial eta squared value is converted into a percentage by multiplying by 100, then the number 70 is obtained. Therefore, it could be argued that 70 percent of the reading posttest is explained by self assessment, while self assessment can explain only 60 percent of the variance of listening posttest. Thus, it could be strongly argued that the impact of self assessment on reading skill is higher than its impact on listening skills.

Such a difference is due to the fact that listening comprehension is different from the reading skill (Nunan 2003). While listening and reading have some similarities, there are two major differences between them. First, listening happens in real time and people listen and have to comprehend immediately without going back, reviewing, and looking up the meaning of unknown words. Second, despite being a receptive skill, it often happens in the midst of a conversation which requires productive and spoken responses (Nunan 2003).

The results also show that self-assessment has a significant effect on improving intermediate language learners listening and speaking skills. Partial Eta Squared of reading and listening tests were .74 and .67, respectively. Therefore, it could be suggested that the impact of self-assessment on the reading skill of intermediate students is higher than its effect on listening skills because self-assessment could explain 74 percent of the variance in the reading posttest; whereas, it could explain 67% of the variance in the listening posttest.

Another interesting finding of the present study is the difference between Partial Eta Squared of posttests of the beginner students and the intermediate language learners. This finding was not found in previous studies. The difference between productive and receptive skills could be an explanation for this finding (Nunan 2003; Harmer 2009) or it could be the result of affective factors and psychological states of beginner and intermediate language learners as well as the rate of their dependence on the teachers. As stated by Richards and Rodgers (2001), beginner language learners are more dependent on the teachers than advanced language learners. Thus, beginner language learners in comparison with intermediate or advanced language learners may not benefit much from self-assessments.

An important implication of this study is that self-assessments are important for educational purposes, therefore teachers must somehow support beginner language learners and provide the necessary feedback while learners are self-assessing their own progress in language skills. The effect of self-assessment on listening and reading might be related to many other factors which need further investigation. The findings of this study are consistent with McDonald and Boud's study (2003) which found that when learners assess their own learning, their learning will be promoted to a higher extent.

The results of the current study are also similar to Butler and Lis' study (2005) whereby young EFL learners exposed to self-assessment had shown improvement in their English performance and in their confidence in learning English. In addition, in line with more recent studies (Black & William 1998; Pellegrino et al. 2001), formative self assessment seems to have a significant positive effect on students' learning.

#### PEDAGOGICAL IMPLICATIONS

This study has implications for language learners, language teachers, curriculum designers and researchers. First, as Barbot (1991) argues, self-assessment is believed to encourage increased sophistication in learner awareness, helps learners to make accurate judgments on their own abilities and to acquire evaluation techniques that comprise the whole learning process; and makes learners see errors as something helpful. It is also seen as a potentially useful tool for teachers since it provides the information on learning styles and on areas needing remediation and feedback. Moreover, language teachers can make use of self-assessment as an alternative to traditional teacher-made tests so that they can filter the affective factors such as test phobia, stress, and anxiety among language learners.

Second, textbook and curriculum designers who provide teaching materials for Iranian English learners should focus on learners' needs and include more examples and varieties of the self-assessment in text books, teaching materials and learning activities, so that the learners have more opportunities for practicing and using self-assessment. Finally, this project can provide more insights and ideas for further research and can help future researchers in their work.

#### CONCLUSION

This study investigated the impact of self-assessment on beginner and intermediate EFL learners' listening and reading skills. Based on the results of this study, it could be concluded that self-assessment had significantly improved language learners' receptive skills. It could be also inferred that self-assessment has a significant effect on both beginner and intermediate language learners' receptive skills. In addition, it could be concluded that beginner language learners need more support and feedback from their teachers than intermediate level language learners. Finally, it is suggested that self-assessment might have a significant impact on other skills as well such as on speaking and writing and may be possibly related to some other variables such as gender, age, personality-type and teachers' attitudes towards self assessment.

#### REFERENCES

- Alderson, C.A. 2005. *Diagnosing Foreign Language Proficiency: The Interface between Learning and Assessment*. New York, NY Continuum.
- Abbasszadeh, S. 2012. The impact of self- assessment on Iranian EFL learners' writing and speaking. Unpublished Master thesis. Yasouj University, Yasouj, Iran.
- AlFallay, I. 2004. The role of some selected psychological and personality traits of the rater in the accuracy of self- and peer-assessment. *System* 32(3): 407-425.
- Bachman, L. F. & Palmer, A.S. 1996. *Language Testing in Practice*. Oxford: Oxford University Press.
- Bailey, K. 1998. Working for washback: A review of the washback concept in language testing. *Language Testing* 13(3): 257-79.
- Barbot, M. J. 1991. New approaches to evaluation in self-access learning (trans. form. French). *Études de Linguistique Appliquée* 79: 77-94.
- Black, P. & Wiliam, D. 1998. Assessment and classroom learning. *Assessment in Higher Education* 5(1): 7-74.
- Blatchford, P. 1997. Students' self assessment of academic attainment: accuracy and stability from 7 to 16 years and influence of domain and social comparison. *Educational Psychology* 17: 354-360.
- Blanche, P. 1990. Using standardized achievement and oral proficiency tests for self-assessment purposes: the DLIFLC study. *Language Testing* 7(2): 202-29.
- Boud, D. 2000. Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education* 22(2): 151-167.
- Brantmeier, C. 2006. Advanced L2 learners and reading placement: self-assessment, computer-based testing, and subsequent performance. *System* 34(1): 15-35.
- Brantmeier, C. 2005a. Nonlinguistic variables in advanced L2 reading: learner's self-assessment and enjoyment. *Foreign Language Annals* 38(4): 493-503.
- Brantmeier, C. 2005b. Effects of reader's knowledge, text type, and test type on L1 and L2 reading comprehension. *The Modern Language Journal* 89(1): 37-53.
- Brown, D. H. 2004. *Language Assessment: Principles and Classroom Practices*. London: Pearson Education, Inc.
- Brown, G., Bull, J. & Pendlebury, M. 1997. *Assessing Student Learning in Higher Education*. Routledge, London.
- Butler, Y. & Li, J. 2003. *The Effect of Self-assessment among Young Learners of English*. University of Pennsylvania.
- Butler, Y.G. & Lee, J. 2010. The effect of self-assessment among young learners of English, *Language Testing* 27: 5-31.
- Calderon, M. 1991. *Promoting Language Proficiency and Academic Achievement Through Cooperation*. ERIC Document, ERIC #: ED436983.
- Carton, F. 1993. Self-evaluation at the heart of learning. *Le Français dans le Monde* (special number): 28-35.
- Cheng, L. & Curtis, A. 2012. Test Impact and Washback: Implications for Teaching and Learning. In *Cambridge Guide to Second Language Assessment*, edited by Coombe, C., Sullivan, B.O., Davidson, P. & Stoyhoff, S. Cambridge: Cambridge University Press.
- Collins, A. 1993. A study of the Provision of Modern Language Pupils with Special Educational Needs. Unpublished M.ED. dissertation, The Queen's University of Belfast.

- Dearing, R. 1997. *Higher Education in the Learning Society*. London: HMSO.
- Glaska, A. & Krekeler, C. 2008. *Self-assessment of Pronunciation*. Surrey University Press.
- Davidson, F. & Henning, G. 1985. A self-rating scale of English difficulty. *Language Testing* 2: 164-169.
- Dickinson, L. 1987. *Self-instruction in Language Learning*. London: Cambridge University Press.
- Falchikov, N. 1997. Why do lecturers involve students in assessment? In: Paper delivered at the 2nd North Umbria Assessment Conference, Encouraging Partnership in Assessing Learning, 3-5 September, University of Northumbria, Newcastle.
- Freeman, R. & Lewis, R. 1998. *Planning and Implementing Assessment*. Kogan Page, London.
- Fallows, S. & Balasubramanyan, C. 2001. Multiple approaches to assessment: reflections on the use of tutor, peer and self-assessment. *Teaching in Higher Education* 6: 229-246.
- Gardner, R. 1985. *Social Psychology and Second Language Learning*. London: Arnold.
- Falchikov, N. & Goldfinch, J. 2000. Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research* 70: 287-322.
- Falchikov, N. & Boud, D. 1989. Student self-assessment in higher education: a meta-analysis. *Review of Educational Research* 59(4): 395-430.
- Graham, S. 2004. Giving up on modern foreign languages? Students' perceptions of learning French. *The Modern Language Journal* 88(2): 171-191.
- Hanrahan, S. & Isaacs, G. 2001. Assessing self- and peer-assessment: the students' views. *Higher Education Research and Development* 20: 53-70.
- Harmer, J. 2009. *The Practice of English Language Teaching*. PLACE: Pearson Longman.
- Heilenman, K. 1991. Self-assessment and placement: a review of the issues. In *Assessing Foreign Language Proficiency of Undergraduates, AAUSC Issues in Language Program Direction*, edited by Teschner, R.V. Heinle & Heinle, Boston.
- Higgins, R., Hartley, P. & Skelton, A. 2001. Getting the message across: the problem of communicating assessment feedback. *Teaching in Higher Education* 6(2): 269-274.
- Ivanic, R., Clark, R. & Rimmershaw, R. 2000. What am I supposed to make of this? The message conveyed to students by tutors' written comments. In *Student Writing in Higher Education: New Contexts*, edited by Lea, M.R., Stierer, B. SHRE/Open University Press, Buckingham.
- Jones, L. & Fletcher, C. 2002. Self-assessment in a selective situation: an evaluation of different measurement approaches. *Journal of Occupational and Organizational Psychology* 75: 145-161.
- Jawah, C., Macfarlane-Dick, D., Matthew, R., Nicol, D., Ross, D. & Smith, B. 2004. Enhancing Student Learning through Effective Formative Feedback. In *The Higher Education Academy*, LTSN, London.
- Janssen-van Dielen, A. 1989. The development of a test of Dutch as a foreign language: the validity of self-assessment by inexperienced subjects. *Language Testing* 6(1): 30-46.
- Kato, F. 2009. Student preferences: Goal-setting and self-assessment activities in a tertiary education environment. *Language Teaching Research* 13(2): 177-199.
- Lejk, M. & Wyvill, M. 2001. Peer assessment of contributions to a group project: a comparison of holistic and category-based approaches. *Assessment and Evaluation in Higher Education* 26: 62-72.
- Li, L. 2005. Some refinement on peer assessment of group projects. *Assessment and Evaluation in Higher Education* 26: 5-18.
- Lindholm-Leary, K. & Borsato, G. 2002. *Impact of Two-Way Immersion on Students' Attitudes toward School and College*. Eric Digest. ERIC Document # ED464541.
- Livesey, D. 1992. *An Application of the Theory of Reasoned Action for Relating Attitude, Social Support and Behavioral Intention in an EFL Setting*. Paper presented at the Annual meeting of the Teachers of English to Speakers of Other Languages 26th, Vancouver, BC, Canada, 3-7 March, 1992.
- Lewkowicz, J.A. & Moon, J. 1985. Evaluation, a way of involving the learner. In *Lancaster Practical Paper in English Language Education* (Vol. 6: Evaluation), edited by Alderson, J.C. Oxford: Pergamon Press.
- Matsuno, S. 2009. Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing* 26(1): 75-100.
- Morton, L., Lemieux, C., Diffey, N. & Awender, M. 1999. Determinants of withdrawal from the bilingual career track when entering high school. *Guidance and Counseling* 14: 1-14.
- McDonald, B. & Boud, D. 2003. The Impact of self-assessment on achievement: the in??? *Education* 10(2): 209-220.
- McDonald, B. & Boud, D. 2003. The impact of self-assessment on achievement: the effects of self-assessment training on performance in external examinations. *Assessment in Education* 10: 209-220.
- Nunan, D. 2003 *Practical English Language Teaching*. Singapore: McGraw Hill.
- Oscarson, M. 1989. Self-assessment of language proficiency: rationale and applications. *Language Testing* 6(1): 1-13.
- Oscarson, M. 1997. Self-assessment of foreign and second language proficiency. In *Language Testing and Assessment* (Vol. 7, 175-87), edited by Clapham, C. & Corson, D. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Oscarson, M. 1989. Self-assessment of language proficiency: Rationale and applications. *Language Testing* 6(1): 1-13.
- Orsmond, P., Merry, S., Reiling, K. (2000). The use of student derived marking criteria in peer and self assessment. *Assessment and Evaluation in Higher Education* 25: 23-38.
- Patri, M. 2002. The influence of peer feedback on self- and peer-assessment of oral skills. *Language Testing* 19: 109-131.
- Pellegrino, J.W., Chudowsky, N. & Glaser, R. 2001. *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academic Press.
- Peirce, B. M., Swain, M. & Hart, D. 1993. Self-assessment, French immersion and locus of control. *Applied Linguistics* 14: 25-42.
- Richards, J.C. & Rodgers, T.S. 2001. *Approaches and Methods in Language Teaching*. Cambridge: Cambridge University Press.
- Shrauger, J.S. & Osberg, T.M. 1981. The relative accuracy of self-predictions and judgments by others of psychological assessment. *Psychological Bulletin* 90: 322-351.



- Stefani, L.J. 1998. Assessment in partnership with learners. *Assessment and Evaluation in Higher Education* 23(4): 339-350.
- Saavedra, R. & Kwun, S. K. 1993. Peer evaluation in self-managing work groups. *Journal of Applied Psychology* 78: 450-462.
- Stefani, L. A. J. 1994. Peer, self, and tutor assessment: relative reliabilities. *Studies in Higher Education* 19: 69-75.
- Sung, Y.T., Chang, K.E., Chiou, S.K. & Hou, H.T. 2005. The design and application of a Web-based self- and peer-assessment system. *Computers and Education: An International Journal* 45: 187-202.
- Sung, Y.T., Chang, K.E., Yu, W.C. & Chang, T.H. 2009. Enhancing teachers' learning and reflection through structured digital portfolios. *Journal of Computer Assisted Learning* 2(1): 21-35.
- Stankowv, L. 1998. Calibration curves, scatter plots, and the distinction between general knowledge and perceptual tests. *Learning and Individual Differences* 8: 28-51.
- Taras, M. 2001. The use of tutor feedback and student self-assessment in summative assessment tasks: toward transparency for students and for tutors. *Assessment and Evaluation in Higher Education* 26: 605-614.
- Taras, M. 2002. Using assessment for learning and learning from assessment. *Assessment and Evaluation in Higher Education* 27: 501-510.
- Teweles, B. 1995. Motivation as a two-sided coin: motivational differences between college-level Chinese and Japanese learners of EFL. *Texas Papers in Foreign Language Education* 2: 1-22.
- Topping, K. 1998. Peer assessment between students in colleges and universities. *Review of Educational Research* 68: 249-276.
- Thomson, C.K. 1996. Self-assessment in self-directed learning: issues of learner diversity. In *Taking control: Autonomy in Language Learning*, edited by Pemberton, R.; Li, E.; Or, W. & Pierson, H. Hong Kong: Hong Kong University Press.
- Todd, R. W. 2002. Using Self-Assessment for Evaluation. *English Teaching Forum* 40(1): 16-19.
- Upshur, J. 1975. Objective evaluation of oral proficiency in the ESOL classroom. In *Paper on Language Testing 1967-1974*, edited by Palmer, L. & Spolsky, B.K. TESOL, Washington, DC.
- Von Elek, T. 1987. A test of Swedish as a second language: an experiment in self-assessment. In *New Directions in Language Testing*, edited by Li, Y., Fok, A., Lord, R., Low, G. Oxford University Press, Oxford.
- Wesche, M., Morrison, F., Ready, D. & Pawley, C. 1990. French immersion: postsecondary consequence for individuals and universities. *Modern Canadian Language Review* 46: 430-451.
- Xiao, Y. & Lucking, R. 2008. The impact of two types of peer assessment on students' performance and satisfaction within a Wiki environment. *The Internet and Higher Education* 11(4): 186-193.

Goudarz Alibakhshi\*  
 Department of English Language Literature  
 Faculty of Foreign Languages and Literature  
 Allameh Tabataba'i University

Hassan Sharakipour  
 Islamic Azad University

Corresponding author\*; email: alibakhshi\_goudarz2000@yahoo.com

Received : 1 September 2013

Accepted : 4 April 2014

