

ANXIETY: A SOURCE OF TEST BIAS?

Parviz Ajideh; Massoud Yaghoubi-Notash; Abdolreza Khalili

Parvizaj@gmail.com; Masoud.yaghoubi@gmail.com; Abdolrezaa.khalili@gmail.com

English Language Department, University of Tabriz, Tabriz, Iran
East Azerbaijan, Tabriz, Iran

Article History:

Received:
March 3, 2017

Revised:
May 24, 2017

Accepted:
July 26, 2017

Corresponding Author:

abdolrezaa.khalili@gmail.com

Abstract: The examination of the various individual learner differences as bias factors in language test performance has been a major concern in the field of language testing for almost three decades. The present study investigated the foreign language classroom anxiety and test anxiety as sources of bias in English vocabulary and grammar tests. Based on this objective, first, 158 intermediate EFL learners were selected as the participants. Second, the participants respectively took the Foreign Language Classroom Anxiety Scale (Horwitz, Horwitz, & Cope, 1986), Test Anxiety Scale (In' nami, 2006), the vocabulary test of the study, and the grammar test of the study in a 2-week period. The standard multiple regression was employed for data analysis. The results revealed that, classroom anxiety and test anxiety had significant negative correlations with the vocabulary and grammar test results. The results of the study may provide certain theoretical implications for testing specialists regarding the redefinition of the construct of second language ability in the process of test validation.

Keywords – Foreign Language Classroom Anxiety, Individual Learner Differences, Test Anxiety, Test Bias, Test Performance

INTRODUCTION

There is considerable variation among language learners regarding their success in language acquisition (Ellis, 2004). This variation is limited to the rate of acquisition for the children who are acquiring their native language. That is, although children differ in the speed of acquiring their mother tongue, they achieve perfect mastery of every aspect of that language (Bley-Vroman, 1988; Clark, 2009). However, this is not true for second

language learners. As Bley-Vroman (1988) noted, most of these learners do not achieve a native-like competence in the use of the second language. According to him:

The general characteristics of foreign language learning tend to the conclusions that the domain-specific language acquisition of children ceases to operate in adults, and in addition, that foreign language acquisition resembles general adult learning in fields for which no domain-specific learning system is believed to exist (p. 25).

Therefore, in the case of second language acquisition, the variation involves both The learners' rate and ultimate level of achievement (Ellis, 2004, 2008). According to Ellis (2004), the differences in achievement among second language learners may stem from three general sets of factors including: *social*, *cognitive*, and *affective* factors. As he further argued, since the cognitive and affective factors lie inside the language learner, the researchers have investigated them as individual learner differences. These differences are "enduring personal characteristics that are assumed to apply to everybody and on which people differ by degree" (Dörnyei, 2005, p. 4). They are "factors specific to individual learners which may account for differences in the rate at which learners learn and their level of attainment" (Richards & Schmidt, 2010, p. 278).

Horwitz (2000) noted that, the investigation of the individual learner differences has always been a major concern in the field of applied linguistics. However, as she argued, there has been an evolutionary and noticeable change regarding the terms that are used to refer to these differences. According to her:

The terms good and bad, intelligent and dull, motivated and unmotivated have given way to a myriad of new terms such as integratively and instrumentally motivated, anxious and comfortable, field independent and field sensitive, auditory and visual (p. 532).

As Ellis (2008) stated, the investigation of individual learner differences has been motivated by different purposes. According to him, some of the studies have tried to identify the language learners who are likely to be more successful in studying certain foreign languages in comparison with the others (e.g. Carrol, 1981). Other studies have tried to determine the relationship between different individual characteristics and second language acquisition (e.g. Glikman, Gardner & Smythe, 1982). Finally, a number of studies have investigated the individual learner differences as potential sources of bias in language learners' test performance (e.g. Hansen & Stanfield, 1981).

According to Bachman (1990), the individuals' scores on different tests may be influenced by both a group of personal characteristics such as cognitive style and ambiguity tolerance, and a number of group characteristics including race and ethnic background. As he further noted, unlike the random factors which have an unpredictable

and transient effect on the learners' scores, the personal or individual characteristics influence the learners' scores regularly. However, as he explained, these characteristics are not part of the language ability that the language tests measure, and as a result, are regarded to be *systematic* sources that influence the validity of the inferences that are made based on the test results. As he stated, the “systematic differences in test performance that are the result of differences in individual characteristics other than the ability being tested” (p. 271) are sources of *test bias*. In other words, a test or a single test item is biased “if its scores are consistently too high or too low, for an individual test taker or a group of test takers” (Richards & Schmidt, 2010, p. 53).

As Bachman (1990) pointed out, the studies of test bias are essential in the field of language testing since they provide a better understanding of the validity of the language tests. According to him, these studies “raise questions about the extent to which language abilities as constructs are independent of the content and context of the language use elicited in their measurement” (p. 279). Moreover, as he explained, these studies may help us judge about the measurement value of the different tests as instruments for testing the language ability. Furthermore, as he noted, they may help us to determine the characteristics of successful language learners and the role of the individual learner differences in the process of language acquisition. Finally, as Farhady (1982) argued, these studies may help us redefine the construct of *language ability*.

A review of the related literature (e.g. Alpert & Haber, 1960; Arnold, 1999; Horwitz, 2001; Kleijn, Van der Ploeg, & Topman, 1994; Scovel, 1978) shows that, among the individual learner differences, anxiety has been investigated by a number of SLA researchers.

Anxiety is one of the affective variables that have received a lot of attention in the field of SLA (Ellis, 2008). It can be defined as “the subjective feeling of tension, apprehension, nervousness, and worry associated with an arousal of the automatic nervous system” (Spielberger, 1983, p. 1). In other words, it is associated with the “feelings of uneasiness, frustration, self-doubt, apprehension or worry” (Scovel, 1978, p. 134). Anxiety is “quite possibly the affective factor that most pervasively obstructs the learning process” (Arnold & Brown, 1999, p. 8). Furthermore, Pramuktiyono and Wardhono (2016) stated that anxiety is merely promising if language educators are well aware of the existence of the anxiety.

There is unanimous approval among various researchers that anxiety affects second language learning and performance (Arnold, 1999; Horwitz & Young 1991; MacIntyre, 1999, 2002; Young, 1999).

However, the investigation of anxiety has always been plagued by theoretical challenges (Oxford, 1999). More specifically, this affective variable cannot be easily categorized with the other affective characteristics such as the motivation or personality factors (Bailey, 1983). Moreover, anxiety “is usually not seen as a unitary factor but a complex made up of constituents that have different characteristics” (Dörnyei, 2005, p. 198).

Anxiety can be *facilitative* or *debilitative* (Alpert & Haber, 1960, Scovel, 1978). The facilitative anxiety is associated with *emotionality* and the debilitating anxiety is related to *worry*. It can be argued that, *worry* is the cognitive component of anxiety and has a negative effect on language performance while *emotionality* is its affective component and may have a positive effect on language learning (Dörnyei, 2005).

Moreover, anxiety may have different types including: *trait*, *state*, and *situation-specific* anxiety (Oxford, 1999). Trait anxiety is “a more permanent predisposition to be anxious” (Scovel, 1978, p. 135) and as a result should be treated as a personality factor (Ganschow & Sparks, 1996). State anxiety is “the apprehension that is experienced at a particular moment in time as a response to a definite situation and therefore is a combination of trait and situation-specific anxiety” (Ellis, 2008, p. 691). In the field of SLA, the researchers have focused on a specific kind of situation-based anxiety which is called *foreignlanguage anxiety* (Horwitz, 2001; Horwitz, Horwitz, & Cope, 1986). This kind of anxiety involves “worry and negative emotional reaction aroused when learning or using a second language” (MacIntyre, 1999, p. 27) According to Horwitz (2001), it is an independent variable which may not be correlated with the other kinds of anxiety. Moreover, a number of researchers (e.g. Brown, 2007; Horwitz et al., 1986; MacIntyre & Gardner, 1989, 1991) have argued that, language anxiety consists of a number of sub-components. According to Brown (2007), this kind of anxiety consists of three components including:

- a) Communication apprehension, arising from learners’ inability to adequately express mature thoughts and ideas;
- b) Fear of negative social evaluation, arising from a learner’s need to make a possible social impression on others;
- c) Test anxiety or apprehension over academic evaluation (p. 162).

Among these components, test anxiety has attracted the researchers’ attention more than the others (e.g. Horwitz, 1986; Joy, 2013). It can be defined as a “special case of general anxiety consisting of phenomenological, physiological, and behavioral responses

related to a fear of failure and to experience of evaluation or testing” (Sieber, 1980, p. 17). As Knox, Schacht and Turner (1993) stated, this type of anxiety plays a very important role in educational contexts since the failure to manage it can result in “failing courses, dropping out of school, a negative self-concept and a low earning potential” (p. 295).

Kleijn, Van der Ploeg, and Topman (1994) discussed three causes of test anxiety. As they explained, the first cause of this type of anxiety is the inappropriate preparation for the examination which stems from the lack of sufficient learning strategies. The second cause of the test anxiety is the focus on the task-irrelevant stimuli during tests which negatively affects test performance. Finally, as they stated, the third cause of this kind of anxiety is the test takers’ wrong beliefs about their readiness for the tests. That is, contrary to the reality, some of the test takers assume that they have prepared for the test in an appropriate way, and when they are not satisfied with their results, they begin to question their test-taking ability.

However, a close examination of the relevant literature shows that, most of the empirical studies of anxiety have tried to determine its sources (e.g. Bailey, 1983; Matsumoto, 1987; Oxford, 1992) and have ignored its role as a source of bias in language tests. Moreover, these studies have only focused on foreign language classroom anxiety (e.g. Chastain; 1975; Kleinmann; 1978) and have not dealt with the test anxiety. Finally, the majority of these studies (e.g. Horwitz; 1986; MacIntyre& Gardner, 1994) have been conducted in second language contexts, and there is not sufficient information about anxiety in foreign language contexts.

In the English as a Foreign Language (EFL) context of Iran, the empirical studies of anxiety have followed a similar trend. More specifically, there is a lack of research regarding the role of anxiety as a source of test bias in the results of the tests of the second language. The present study was an attempt to deal with the mentioned gaps of the literature regarding the language anxiety. Based on this aim, it tried to provide more information regarding the role of Iranian intermediate-level male EFL learners’ foreign language classroom anxiety and test anxiety as bias factors in their performance on the vocabulary and grammar tests.

METHOD

Design of the study

As Creswell (2011) pointed out, the correlational research design takes two main forms including; the *explanatory design* and the *prediction design*. In explaining the prediction design he stated that:

The purpose of the prediction design is to identify variables that will predict an outcome or criterion. In this form of research the researcher identifies one or more *predictor* variables and a *criterion* or *outcome* variable. A predictor variable is a variable which is used to make a forecast about an outcome in correlational research....The outcome being predicted in correlational research, however, is called the criterion variable (p. 341).

An examination of the purpose, data collection, and data analysis of the present study shows that, it employed a quantitative approach and was conducted based on a predictive correlational design in which the foreign language classroom anxiety and test anxiety were the predictor variables and the learners' performances on the vocabulary and grammar tests were the criterion variables.

Participants

In the present study, 158 intermediate EFL learners were selected from among 324 language learners of a private language institute in Urmia (Iran) as the participants of the study based on their results on the Objective Placement Test (Lesley, Hansen, &Zukowski, 2003). The selected participants: were male, ranged in age from 15 to 26, and had 2 to 3 years of language studies in the language institute. They were from Urmia and were native speakers of Azeri. In order to select these participants, first, the researchers determined the mean value of the 324 language learners' results on the proficiency test of the study. Second, they selected the learners whose score were within 1 Standard Deviation (SD) below and above the mean value of the group.

The instruments and materials of the study

The following instruments and materials were employed in the present study:

Proficiency test

The determination of the proficiency level and the homogeneity of the selected participants are essential in order to guarantee the validity of the inferences that are made based on the results of the empirical studies in the field of second language acquisition (Mackey &Gass, 2016). The present study tried to determine the relationship between the intermediate EFL learners' anxiety and their test performance. Based on this aim, the Objective Placement Test, from New Interchange Passages Placement and Evaluation Package (Lesley, Hansen, &Zukowski, 2003) was employed in order to select the participants of the study. This test consisted of four parts: Listening, Grammar, Vocabulary, and Reading. The Listening section involved 20 recorded items. The Grammar section had

30 items. The Vocabulary section consisted of 30 items and the Reading section had 20 items.

The foreign language classroom anxiety scale

In order to assess the participants' language anxiety, the Foreign Language Classroom Anxiety Scale (FLCAS) (Horwitz et al., 1986) was employed in this study. This questionnaire involves 33 items which are rated on a 5-point Likert scale. The answers to each item range from *strongly agree* to *strongly disagree*. As the researchers stated, in determining the learners' anxiety level a score of 5 is given to the *strongly agree* and a score of 1 is given to *strongly disagree*. According to them, a higher score on FLCAS shows a higher level of language anxiety. In order to determine the reliability of this questionnaire, Horwitz et al. (1986) employed a test-retest method. According to them, since the "test-retest reliability over eight weeks yielded an $r = .83$ ($p < .001$)" it was argued that, the questionnaire was a highly reliable instrument in determining foreign language anxiety. On the other hand, Aida's (1994) factor analysis of the items of this questionnaire showed that, they are highly valid in examining the learners' level of language anxiety.

The test anxiety scale

In order to measure the participants' test anxiety, the modified version of Test Anxiety Scale (TAS) (In'nami, 2006) was employed in this study. The original form of this questionnaire (with 37 yes/no items) was developed by Sarason (1975), and was employed in order to examine the learners' anxiety in taking different kinds of language tests. In an attempt to increase the accuracy of the responses to the questionnaire items, In'nami (2006) changed the yes/no answers to the items with a 5-point Likert scale (i.e. 1= completely disagree; 2 = disagree; 3= neutral; 4= agree; 5= completely agree). Moreover, he factor analyzed the items of questionnaire and stated that, the questionnaire is a highly valid and reliable instrument for determining test anxiety. The employed version of this questionnaire in the present study involves 37 items which are scored on the mentioned 5-point Likert scale. According to In'nami (2006), a high score on this questionnaire indicates a high level of test anxiety.

The vocabulary test

Based on the aims of the present study, a 40-item researcher-made multiple-choice vocabulary test was employed in order to determine the participants' vocabulary test performance. The items of this test were developed based on the vocabulary items of the reading texts of *Intermediate Select Readings* (Lee & Gundersen, 2011). In order to

guarantee that the test was a reliable and valid instrument for measuring the participants' vocabulary knowledge, it was piloted with 75 male ELF learners which had similar characteristics to the participants of the study. Since the items of the test were based on the reading texts that were specifically developed for the intermediate-level language learners, their content validity was guaranteed. However, in order to determine the empirical (concurrent) validity of the test, the results of the selected 75 learners on this test were correlated with their results on the vocabulary section of the Objective Placement Test, (Lesley, Hansen, &Zukowski 2003). The results of the analysis showed that, the empirical validity index of the test was .82 which, as Harris (1969) stated, is regarded to be satisfactory for researcher/teacher-made tests. Moreover, a test-retest method was employed in order to determine the reliability of the test. Based on this aim, the selected 75 EFL learners took this test twice during a one month period and their results on the two sessions were correlated. The results of this analysis showed that, the reliability index of the vocabulary test was .87 which, as Harris (1969) stated, is regarded to be satisfactory for researcher/teacher-made tests.

The grammar test

In order to determine the selected participants' grammar test performance, a 40-item researcher-made multiple-choice grammar test was employed in the present study. Similar to the vocabulary test, the items of this test were based on the reading texts of *Intermediate Select Readings* (Lee &Gundersen, 2011). That is, the researchers extracted the grammar points of these reading texts and developed the test items based on these points. In order to guarantee the reliability and validity of this test, the researchers piloted it with 75 male EFL learners with similar characteristics to selected participants. Since the test items were based on intermediate-level reading texts (i.e. texts of *Intermediate Select Readings*) their content validity was guaranteed. However, in order to determine the empirical (concurrent) validity of the test, the results of the selected 75 learners on this test were correlated with their results on the grammar section of the Objective Placement Test, (Lesley, Hansen, &Zukowski, 2003). The results of the analysis showed that, the empirical validity index of the test was .78 which, as Harris (1969) stated, is regarded to be satisfactory for researcher/teacher-made tests. Moreover, a test-retest method was employed for determining the reliability of the test items. That is, the selected learners took the test twice during a one month period and their results were correlated. Based on the results, the reliability index of the grammar test was .84 which, as Harris (1969) stated, is regarded to be satisfactory for researcher/teacher-made tests.

The procedure of the study

According to Pallant (2007), in order to be able to generalize the results of a study to other relevant studies, there is a need for an appropriate sample size in Multiple Regression. As she further discussed, the following formula may be employed for the calculation of the required sample size: “ $N > 50 + 8m$ (where m shows the number of independent variables)” (p. 148). The present study had 2 independent variables (i.e. classroom anxiety & test anxiety). Therefore, the number of the participants had to be more than 66. However, the researchers decided to select all of the participants with the required characteristics from among the available learners to increase the generalizability of the study results. Based on this issue, in this study, first, 158 intermediate EFL learners were selected from among 324 language learners of a private language institute in Urmia (Iran) as the participants of the study based on their results on the Objective Placement Test (Lesley, Hansen, & Zukowski, 2003). Second, the FLCAS (Horwitz et al., 1986) was administered to the selected participants of the study in order to assess their foreign language classroom anxiety. It took the participants about 15 minutes to answer the items of this questionnaire. Third, the participants received the TAS (In’ nami, 2006) for the assessment of their test anxiety. They answered the items of this questionnaire in about 20 minutes. Fourth, the vocabulary test of the study was administered to the participants for the determination of their English vocabulary test performance. This test took about 45 minutes of the class time. Finally, the participants took the grammar test of the study for the determination of their English grammar test performance. They answered the items of this test in about 45 minutes. The questionnaires and the tests of the study were administered to the participants during 4 sessions in a 2-week period. The researchers employed the Statistical Package for the Social Sciences (SPSS) Version 20 for the data analysis of the study.

FINDINGS AND DISCUSSIONS

Findings

The first research question of the study tried to determine the relationship between the EFL learners’ anxiety and their vocabulary test performance. Based on the aims of this research question, a Standard Multiple Regression test was run between the participant’s results on the language anxiety inventory and test anxiety inventory of the study and their performance on the vocabulary test. In the regression analysis, first, the assumption of multicollinearity had to be checked. In order to check this assumption, the collinearity

diagnostics including *Tolerance* and *Variance Inflation Factor* (VIF) were determined.

According to Pallant (2007):

Tolerance is an indicator of how much of the variability of the specified independent is not explained by the other independent variables in the model. If this value is very small (less than .10), it indicates that the multiple correlation with other variables is high, suggesting the possibility of multicollinearity. The other value given is the VIP, which is just the inverse of the Tolerance value (1 divided by Tolerance). VIF values above 10 would be a concern, indicating multicollinearity (p. 156).

The Tolerance and VIF values of the regression model for the vocabulary test are provided in Table 1 below:

Table 1. The collinearity diagnostics of the learners' anxiety types and vocabulary test performance

Model	Tolerance	VIF
Classroom Anxiety	.999	1.001
Test Anxiety	.999	1.001

As Table 1 shows, the Tolerance values of the model were more than 0.10, and the VIF values were less than 10. Therefore, the multicollinearity assumption was not violated. Moreover, in order to determine the outliers, the Mahalanobis distance value was checked. As Pallant (2007) noted, for a model with 2 independent variables this value should not exceed "13.82" (p. 157). The results of *residuals statistics* for this model are provided in Table 2 below:

Table 2. The residuals statistics of the regression model of the learners' anxiety types and vocabulary test performance

	Minimum	Maximum	Mean	Std. Deviation	N
Mahal. Distance	.070	7.949	1.987	1.803	158
Cook's Distance	.000	.055	.006	.009	158

As Table 2 shows, the maximum value of the Mahalanobis distance (7.949) was less than 13.82, and therefore the outlier assumption was not violated. Finally, in order to

check the remaining assumptions, the value of the Cook's distance was determined. As Pallant (2007) argued, this value should be less than 1. According to Table 2, the maximum value for the Cook's distance (.055) was less than 1, and therefore none of the other assumptions was violated.

Since all of the assumptions of the Multiple Regression were present, the regression model of the learners' anxiety types and vocabulary test performance was evaluated. Table 3 below provides the summary of this model:

Table 3. The regression model summary of the learners' anxiety types and vocabulary test performance

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.394	.155	.144	7.141

According to Table 3, this model explains 0.155 (i.e. *R Square* value) of the variance of the learners' performance on the vocabulary test. That is, this model explains 15.5 percent (*R Square* value multiplied by 100, by shifting the decimal point two places to the right) of the variance in the vocabulary test performance. However, in order to check the statistical significance of the predictive power of the model the results of the ANOVA test of the model had to be checked. The results of this test are provided in Table 4 below:

Table 4. The ANOVA test of the regression model of the learners' anxiety types and vocabulary test performance

Model	Sum of Squares	Df	Mean Square	F	Sig.
Regression	1430.542	2	715.271	14.026	.000
Residual	7802.151	153	50.994		
Total	9232.692	155			

As Table 4 shows, the predictive power of the model was not equal to 0 since the p-value of the ANOVA test .000 (marked as *Sig.*) was less than the level of significance .05.

Finally, in order to determine the contribution of each of the independent variables to the prediction of the variance of the vocabulary test results the *Standardized Coefficients* had to be checked. These results are provided in Table 5 below:

Table 5. The coefficients of the regression model of the learners' anxiety types and vocabulary test performance

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error			
(Constant)	33.757	2.464		13.698	.000
Classroom Anxiety	-.041	.016	-.194	-2.606	.010
Test Anxiety	-.048	.010	-.349	-4.695	.000

According to Table 5, the Beta value for the Test Anxiety variable (-.349) is larger than the other variable. Therefore, it can be argued that, this variable makes a stronger unique contribution to explaining the results of the vocabulary test when the variance explained by the other variable in the model is controlled. Moreover, since the p-value for this variable .000 (marked as *Sig.*) was less than the level of significance .05, it was argued that, this variable made a statistically significant unique contribution to the prediction of the vocabulary test results. Furthermore, Classroom Anxiety (Beta=-.194, Sig=.010) also made a statistically significant contribution to the results of this test. The significant contributions of these variables to the explanation of the results of this test are respectively depicted in Figure 1 and Figure 2 below:

Figure 1. The correlation between the learners' test anxiety and vocabulary test performance

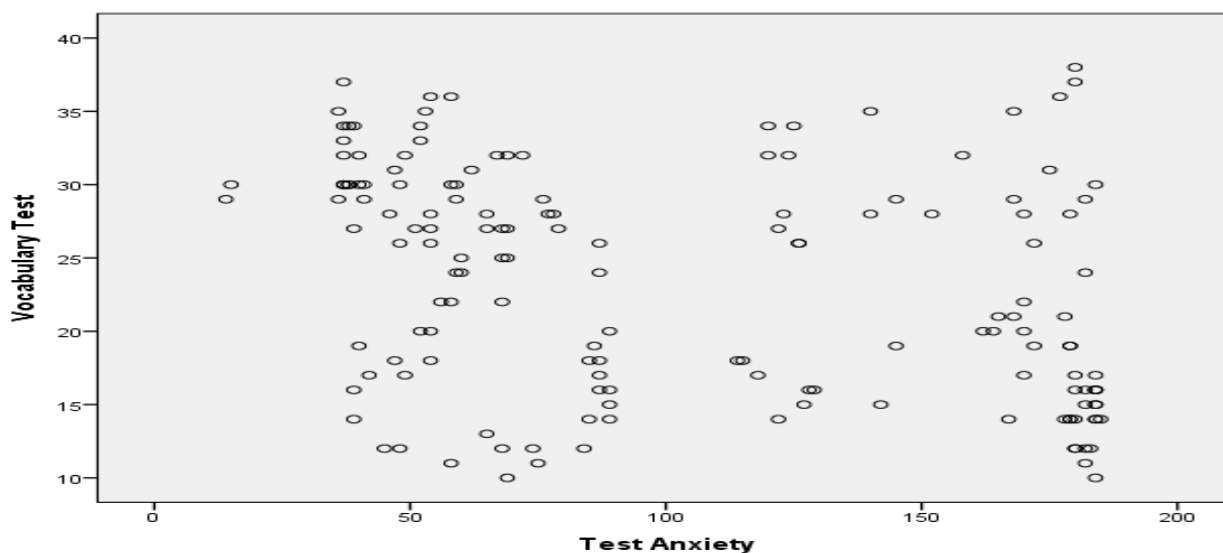
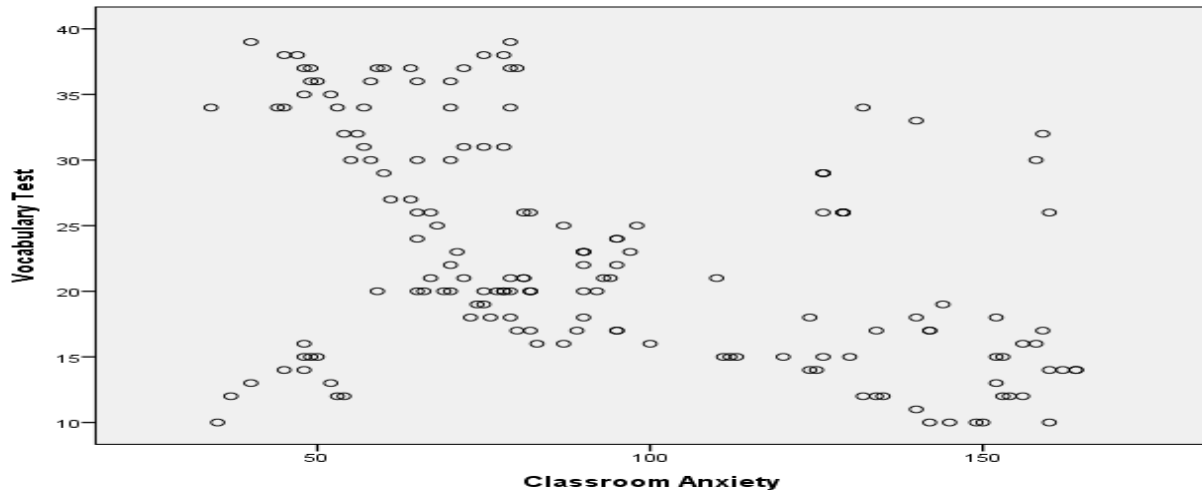


Figure 2. The correlation between the learners' classroom anxiety and vocabulary test performance



The second research question of the study tried to determine the relationship between the EFL learners' anxiety and their grammar test performance. Based on the aims of this research question, a Standard Multiple Regression test was run between the participant's results on the language anxiety inventory and test anxiety inventory of the study and their performance on the grammar test. In the regression analysis, first, the assumption of multicollinearity had to be checked. The Tolerance and VIF values of the regression model for the grammar test are provided in Table 6 below:

Table 6. The collinearity diagnostics of the learners' anxiety types and grammar test performance

Model	Tolerance	VIF
Classroom Anxiety	.957	1.045
Test Anxiety	.957	1.045

As Table 6 shows, the Tolerance values of the model were more than 0.10, and the VIF values were less than 10. Therefore, the multicollinearity assumption was not violated.

Moreover, in order to determine the outliers, the Mahalanobis distance value was checked. The results of *residuals statistics* for this model are provided in Table 7 below:

Table 7: The residuals statistics of the regression model of the learners' anxiety types and grammar test performance

	Minimum	Maximum	Mean	Std. Deviation	N
Mahal. Distance	.112	7.579	1.987	2.913	158
Cook's Distance	.000	.292	.010	.031	158

As Table 7 shows, the maximum value of the Mahalanobis distance (7.579) was less than 13.82, and therefore the outlier assumption was not violated. Moreover, the maximum value for the Cook's distance (.292) was less than 1. Therefore, none of the assumptions was violated. Since all of the assumptions of the Multiple Regression were present, the regression model of the learners' anxiety types and grammar test performance was evaluated. Table 8 below provides the summary of this model:

Table 8: The regression model summary of the learners' anxiety types and grammar test performance

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.328	.107	.096	7.521

According to Table 8, this model explains 0.107 (i.e. *R Square* value) of the variance of the learners' performance on the grammar test. That is, this model explains 10.7 percent (*R Square* value multiplied by 100, by shifting the decimal point two places to the right) of the variance in the grammar test results. However, in order to check the statistical significance of the predictive power of the model the results of the ANOVA test of the model had to be checked. The results of this test are provided in Table 9 below:

Table 9. The ANOVA test of the regression model of the learners' anxiety types and grammar test performance

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	1053.630	2	526.815	9.313	.000
Residual	8767.820	155	56.567		
Total	9821.449	157			

As Table 9 shows, the predictive power of the model was not equal to 0 since the p-value of the ANOVA test .000 (marked as *Sig.*) was less than the level of significance .05.

Finally, in order to determine the contribution of each of the independent variables to the prediction of the variance of the grammar test results the *Standardized Coefficients* had to be checked. These results are provided in Table 10 below:

Table 10. The coefficients of the regression model of the learners' anxiety types and grammar test performance

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error			
(Constant)	19.429	4.240		4.582	.000
Classroom Anxiety	.061	.026	-.180	-2.315	.022
Test Anxiety	-.034	.011	-.239	-3.084	.002

According to Table 10, the Beta value for the Test Anxiety variable (-.239) is larger than the other variable. Therefore, it can be argued that, this variable makes a stronger unique contribution to explaining the results of the grammar test when the variance explained by the other variable in the model is controlled. Moreover, since the p-value for this variable .002 (marked as *Sig.*) was less than the level of significance .05, it was argued that this variable made a statistically significant unique contribution to the prediction of the

grammar test results. Furthermore, Classroom Anxiety (Beta=-.180, Sig=.022) also made a statistically significant contribution to the results of this test. The significant contributions of these variables to the explanation of the results of this test are respectively depicted in Figure 3 and Figure 4 below:

Figure 3. The correlation between the learners' test anxiety and grammar test performance

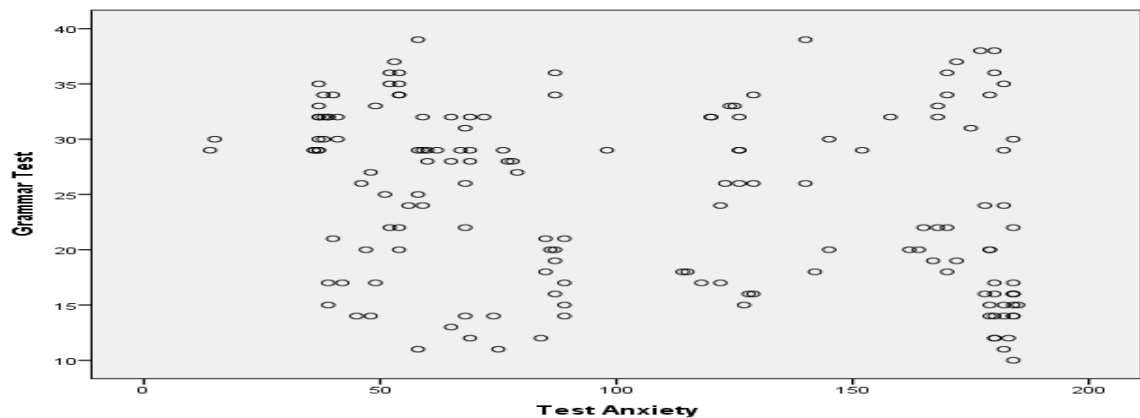
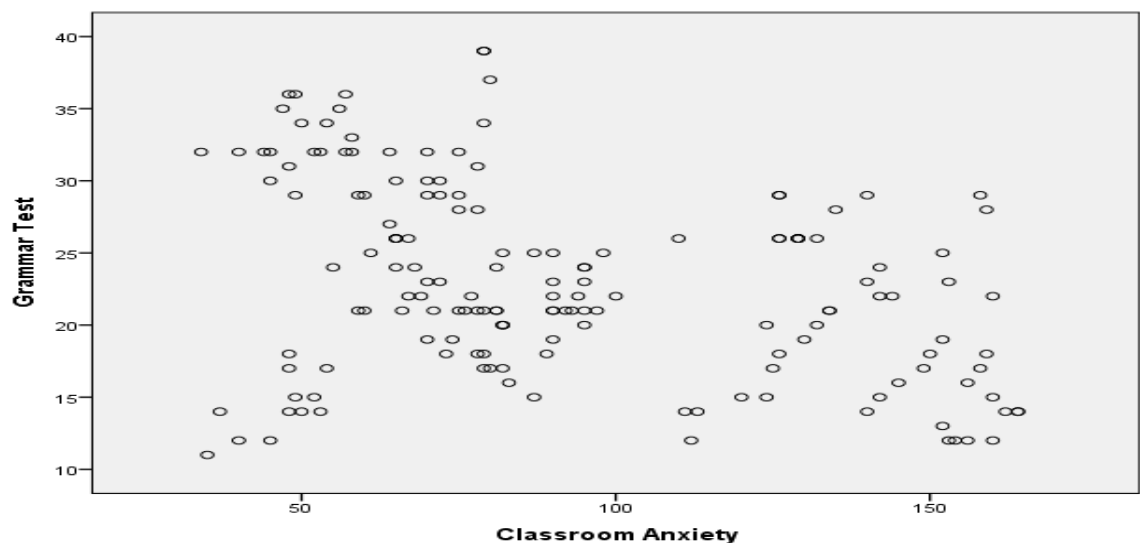


Figure 4. The correlation between the learners' classroom anxiety and grammar test performance



Discussions

The first research questions of the study tried to determine the relationship between the learners' anxiety and their performance on the vocabulary test of the study. More specifically, it tried to determine how much of the variance in the learners' results on the vocabulary test can be explained by the learners' anxiety. The results of the data analysis revealed that, the learners' Test Anxiety and Classroom Anxiety significantly contributed to the explanation of the variance in the results of this test. Moreover, based on the results, Test Anxiety made a stronger contribution to the results in comparison with the Classroom Anxiety. These results are in line with the results of the studies by Horwitz

(1986), MacIntyre and Gardner (1991, 1994) who have reported significant negative correlations between the language anxiety types and second language test performance.

According to MacIntyre and Gardner (1991), the second language anxiety including both the classroom anxiety and the test anxiety can have debilitating effects on the language learners' performance on second language tests. As they explained, the beginner-level learners usually do not experience high levels of language anxiety. That is, anxiety may not be a powerful predictor of language achievement at the lower proficiency levels. However, as they noted, when the learners reach the post-beginner and the intermediate levels, they may experience higher levels of language anxiety as a result of unsatisfactory learning experiences in their classrooms. According to them, this kind of anxiety may prevent the learners from focusing and learning the different aspects of the second language. As they concluded, at this stage, the increase of the second language anxiety acts as a predictor of language achievement in general, and second language test performance in particular. More specifically, the higher levels of language anxiety lead to the learners' weak performance on second language tests.

Based on these issues, it can be argued that, the intermediate-level learners of the present study with high levels of classroom and test anxiety had a weaker performance on the vocabulary test of the study in comparison with the others since they were not able to focus and learn the formal aspects of the second language including its vocabulary. Therefore, it can be concluded that, the learners' test anxiety and classroom anxiety may be systematic sources of test bias and affect their performance on the vocabulary tests of the second language.

The second research question of the study tried to determine the relationship between the learners' anxiety and their performance on the grammar test of the study. The results of the data analysis revealed that, the learners' Test Anxiety and Classroom Anxiety significantly contributed to the explanation of the variance in the results of this test. Moreover, based on the results, Test Anxiety made a stronger contribution to the results in comparison with the Classroom Anxiety. These results are in line with the results of the studies by Birjandi and Alemi (2010), and Salehi and Marefat (2014) who have reported significant negative correlations between the anxiety types and second language test performance.

As MacIntyre and Gardner (1991) stated, the students' language anxiety usually has a negative effect on their second language acquisition. As they explained, the learners at the intermediate proficiency levels of language acquisition are likely to be more affectively inhibited in comparison with the beginner-level learners. According to them, most of these

learners experience a higher level of language anxiety in the language classrooms in comparison with the learners who are at the basic or elementary proficiency levels. As they concluded, the learners with this kind of anxiety are not able to properly pay attention to and learn the different aspects of the second language including the grammatical structures.

Based on these issues, it can be argued that, the intermediate-level learners of the present study with high levels of classroom and test anxiety had a weaker performance on the grammar test of the present study in comparison with the others since they were not able to recognize and pay attention to the grammatical structures of the second language and could not use them accurately. Therefore, it can be concluded that, the learners' test anxiety and classroom anxiety may be systematic sources of test bias and affect their performance on the grammar tests of the second language.

Finally, it should be noted that, the results of the present study do not support the results of the studies by Chastain (1975) who reported mixed results regarding the relationship between language anxiety and test performance. Moreover, the results are in contrast with the results of the study by Kleinmann (1978) who reported a significant positive correlation between the language anxiety and second language test performance. Finally, the results are not in line with the results of the study by In'nami (2006) who did not find any significant relationship between the participants' test anxiety and their second language test performance.

According to Bailey (1983), the learners' second language anxiety may stem from: their competitive nature in classrooms, their perceived relationship with the second language teacher, and the nature of the evaluation system of the classroom. Moreover, as Oxford (1992) noted, this kind of anxiety may be related to the learners' fear of losing themselves in the target culture. Finally, as Gregersen and Horwitz (2002) stated, the language anxiety may be influenced by the language learners' personality traits such as perfectionism in language acquisition.

According to Skehan (1989), the different sources of anxiety may influence the relationship between language anxiety and second language acquisition. That is, the various factors that provoke anxiety may modify the relationship between the learners' experienced language anxiety and their performance in the second language.

Based on these issues, it can be argued that, the participants' affective variables (e.g. their competitive nature in classrooms, their fear of losing themselves in the target culture, & their personality traits) along with the situational factors (e.g. the nature of the evaluation system of the classroom) have led to the difference between the results of the present study and the studies by Chastain (1975), Kleinmann (1978), and In'nami (2006).

CONCLUSIONS AND SUGGESTION

The present study investigated the relationship between the EFL learners' anxiety and their vocabulary and grammar test performance. The results of the study showed that, the learners' test anxiety and classroom anxiety had significant negative correlations with their performances on the vocabulary and grammar tests of the study. Based on these results, the EFL teachers are recommended to employ reliable and valid instruments (e.g. FLCAS& TAS) in order to determine their learners' anxiety types. Moreover, the teachers are recommended to reduce their learners' second language anxiety by the means of various techniques. For example, they can motivate the learners to participate in the activities of the classroom and can encourage them to express their opinions about the usefulness of the various classroom tasks. Furthermore, they can reduce the learners' anxiety by asking them: to participate in different kinds of pair and group activities, and to employ affective strategies when they are using the second language in the classroom. In addition, the teachers can discuss the problems which most of the learners encounter in the process of language acquisition.

Finally, the language teachers can reduce their learners' test anxiety by explaining the evaluation system of the classroom to them and assuring them that they can pass their tests easily. That is, during the second language course, they should tell their students about: the parts of the lessons which will be included in the language tests, the structure of the tests, and the sessions in which they will take the tests.

Based on the mentioned results of the study regarding the learners' anxiety types, the EFL syllabus designers are recommended to include sufficient pair and group activities in the EFL textbooks in order to reduce the learners' second language anxiety. Moreover, they should include certain sections in the textbooks in which the learners become familiar with various kinds of test items and are instructed to employ appropriate test-taking strategies in order to answer them. Finally, they are recommended to include sufficient information regarding the affective strategies in the textbooks in order to reduce the language learners' anxiety types.

Finally, as Skehan (1989) noted, the language testing specialists are recommended to adopt a research-then-theory approach in the studies of individual learner differences in order to provide more information regarding the random, non-linear, and context-specific role of these differences in the explanation of the variance in the results of different measures of the second language.

However, it should be noted that, there is a need for various empirical studies of individual learner differences in different learning contexts and educational settings in order to make wide-reaching conclusions about the role of these differences as test bias factors. For instance, the future studies should investigate larger samples including both male and female second language learners. Moreover, they should involve language learners from different age groups. The investigation of these personal attributes may help to answer certain questions regarding the differential development of language ability based on the learners' age and gender (Bachman, 1990). Furthermore, the future studies should involve language learners from different mother tongues, and language proficiency levels in order to provide more information regarding the non-linear and variable role of the individual learner differences in the explanation of the variance in second language tests.

REFERENCES

- Aida, Y. (1994). Examination of Horwitz, Horwitz and Cope's construct of foreign language anxiety: The case of students of Japanese. *The Modern Language Journal*. 78 (2), 155-68.
- Alpert, R., & Haber, R. N. (1960). Anxiety in academic achievement situations. *Journal of Abnormal and Social psychology*, 61(10), 207-215.
- Arnold, J. (1999). *Affective language learning*. Cambridge: Cambridge University Press.
- Arnold, J., & Brown, H. D. (1999). A map of the terrain. In J. Arnold (Ed.), *Affect in language learning* (pp. 1-24). Cambridge: Cambridge University Press.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bailey, K. (1983). Competitiveness and anxiety in adult second language learning: Looking at and through the diary studies. In H. Seliger & M. Long (Eds.), *Classroom-oriented research in second language acquisition* (pp. 67-102). Rowley: Newbury House.
- Birjandi, P., & Alemis, M. (2010). The impact of test anxiety on test performance among Iranian EFL learners. *Broad Research in Artificial Intelligence and Neuroscience*, 1(4), 44-58.
- Bley-Vroman, R. (1988). The fundamental character of foreign language learning. In W. Rutherford & M. Sharwood Smith (Eds.), *Grammar and second language teaching: A book of readings* (pp. 19-30). New York: Newbury House.

- Brown, H. D. (2007). *Principles of language learning and teaching* (5th Ed.). New York: Pearson Education.
- Carroll, J. B. (1981). Twenty-five years of research in foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83-118). Rowley: Newbury House.
- Chastain, K. (1975). Affective and ability factors in second language acquisition. *Language Learning*, 25(1), 153-161.
- Clark, E. V. (2009). *First language acquisition* (2nd Ed.). Cambridge: Cambridge University Press.
- Creswell, J. W. (2011). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (4th Ed.). New York: Pearson Education.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah: Lawrence Erlbaum.
- Ellis, R. (2004). Individual differences in second language learning. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 525–551). Oxford: Blackwell.
- Ellis, R. (2008). *The study of second language acquisition* (2nd Ed.). Oxford: Oxford University Press.
- Farhady, H. (1982). Measures of language proficiency from the learner's perspective. *TESOL Quarterly*, 16(1), 43-59.
- Ganschow, L., & Sparks, R. (1996). Anxiety about foreign language learning among high school women. *The Modern Language Journal*, 80(2), 199-212.
- Glikzman, L., Gardner, R. C., Smythe, P. (1982). The role of integrative motivation on students' participation in the French classroom. *Canadian Modern Language Review*, 38(3), 625-647.
- Gregersen, T., & Horwitz, E. K. (2002). Language learning and perfectionism: Anxious and non-anxious language learners' reaction to their own oral performance. *The Modern Language Journal*, 86(4), 562-570.
- Hansen, J., & Stansfield, C. (1981). The relationship between field dependent-independent cognitive styles and foreign language achievement. *Language Learning*, 31(2), 349-367.
- Harris, D. P. (1969). *Testing English as a second language*. New York: McGraw-Hill.
- Horwitz, E. K. (1986). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. *TESOL Quarterly*, 20(3), 559-562.
- Horwitz, E. K. (2000). Teachers and students, students and teachers: An ever-evolving partnership. *The Modern Language Journal*, 84(4), 523-535.

- Horwitz, E. K. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics*, 21(2), 112–126.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal*, 70(2), 125–132.
- Horwitz, E. K., & Young, D. (1991). *Language learning anxiety: From theory and research to classroom implications*. Englewood Cliffs: Prentice Hall.
- In'nami, Y. (2006). The effects of test anxiety on listening test performance. *System*, 34 (3), 317-340.
- Joy, J. L. (2013). The altitude of test anxiety among second language learners. *Language Testing in Asia*. 3(4), 1-10.
- Kleijn, W. C., Van der Ploeg, H. M., Topman, R. M. (1994). Cognition, study habits, test anxiety, and academic performance. *Psychological Reports*, 75(3), 1219-1226.
- Kleinmann, H. (1978). The strategy of avoidance in adult second language acquisition. In W. Ritchie (Ed.), *Second language acquisition research* (pp. 157-174). New York: Academic Press.
- Knox, D., Schacht, C., Turner, J. (1993). Virtual reality: A proposal for treating test anxiety in college students. *College Student Journal*, 27(3), 294-296.
- Lee, L., & Gundersen, E. (2011). *Intermediate select readings*. Oxford: Oxford University Press.
- Lesly, T., Hansen, Ch., & Zukowski, J. (2003). *New interchange passages placement and evaluation package*. New York: Cambridge University Press.
- MacIntyre, P. D. (1999). Language anxiety: A review of the research for language teachers. In D. J. Young (Ed.), *Affect in foreign language and second language learning* (pp. 24–45). Boston: McGraw-Hill.
- MacIntyre, P. D. (2002). Motivation, anxiety and emotion in second language acquisition. In P. Robinson (Ed.), *Individual differences in second language acquisition* (pp. 45–68). Amsterdam: John Benjamins.
- MacIntyre, P. D., & Gardner, R. (1989). Anxiety and second language learning: Toward a theoretical clarification. *Language Learning*, 39(2), 251-275.
- MacIntyre, P. D., & Gardner, R. C. (1991). Language anxiety: Its relation to other anxieties and to processing in native and second languages. *Language Learning*, 41(4), 513–534.
- MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, 44(2), 283–305.
- Mackey, A., & Gass, S. (2016). *Second language research: Methodology and design* (2nd Ed.). New York: Tylor and Francis.

- Matsumoto, K. (1987). Factors involved in the L2 learning process. *Japan Association for Language Teaching Journal*, 11(2), 167-192.
- Oxford, R. L. (1992). Who are our students? A synthesis of foreign and second language research on individual differences with implications for instructional practice. *TESL Canada Journal*, 9(2), 30-49.
- Oxford, R. L. (1999). Anxiety and the language learner: New insights. In J. Arnold (Ed.), *Affect in language learning* (pp. 58–67). Cambridge: Cambridge University Press.
- Pallant, J. (2007). *SPSS survival manual: A step by step guide to data analysis using SPSS for windows* (3rd Ed.). Berkshire: McGraw-Hill House.
- Pramuktiyono, A., & Wardhono, A. (2016). Foreign Language Anxiety and Beliefs of EFL Learners At Indonesia. *Indonesian EFL Journal: Journal of ELT, Linguistics, and Literature*, 2(1):14-36.
- Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics* (4th Ed.). Edinburg: Pearson Education.
- Salehi, M., & Marefat, F. (2014). The effects of foreign language anxiety and test anxiety on foreign language test performance. *Theory and Practice in Language Studies*, 4(5), 931-940.
- Sarason, I. G. (1975). The test anxiety scale: Concept and research. In I. G. Sarason & C. D. Spielberger (Eds.), *Stress and anxiety* (pp. 193-217). Washington: Hemisphere.
- Scovel, T. (1978). The effect of affect on foreign language learning: A review of the anxiety research. *Language Learning*, 28(1), 129-142.
- Sieber, J. E. (1980). Defining test anxiety: Problems and approaches. In I. G. Sarason (Ed.), *Test anxiety: Theory, research and applications* (pp. 15-42). Hillsdale: Lawrence Erlbaum Associates.
- Skehan, P. (1989). *Individual differences in second language learning*. London: Edward Arnold.
- Spielberger, C. (1983). *Manual for the state-trait anxiety inventory*. Palo: Consulting Psychologists Press.
- Young, D. J. (1999). *Affect in foreign language and second language learning*. Boston: McGraw-Hill.

Contributors' Biodata

Parviz Ajideh (PhD) is a professor of ELT in the English Department at the University of Tabriz in the Islamic Republic of Iran. He has taught various courses at the University of

Tabriz for 30 years and has authored numerous papers in international journals. His research interests include reading, testing, and translation.

Mssoud Yaghoubi-Notash (PhD) is an assistant professor of ELT and a member of the English Department at the University of Tabriz, Tabriz, Iran. He has authored many papers in international journals and presented papers in national and overseas conferences. His areas of interest include gender and discourse in ELT, task-based language teaching and learning, pragmatics, and EAP.

Abdolreza Khalili is a PhD Candidate in ELT at the University of Tabriz, Tabriz, Iran. He has worked as an EFL instructor at the University of Tabriz for 6 years. Moreover, has taught general English courses at Iran Language Institute for 7 years. He has published various articles in international journals. His areas of interest include language testing, intercultural pragmatics, form-focused instruction, and English for Specific Purposes.