The Effect of Dynamic Assessment Models on L2 Listening and Speaking Anxiety

Abbas Ali Zarei[*][1]

Elham Shishegarha[2]

Abstract

Reducing anxiety in foreign language learning has long been a concern for many teachers. This study focused on exploring the effects of three dynamic assessment models on L2 speaking and listening anxiety. The participants were 120 pre-intermediate Iranian learners of English at a language institute in Qazvin, Iran. The learners were randomly assigned to four groups (three experimental groups and one control group). Before the treatment, the students' homogeneity was checked using Oxford Placement Test (OPT). Then, all the groups were given listening and speaking anxiety questionnaires as pretests. During 10 sessions, the first group received listening and speaking instruction using Buddof's Learning Potential Measurement Approach (LPM); the second group was treated with Guthke's Lerntest Approach; the third group was treated with Testing-the-Limits Approach. Lastly, the control group was taught conventionally in a teacher-fronted way. The same questionnaires were given to the participants in the twelfth session as posttests. Data were analyzed using two one-way analysis of covariance procedures. Significant differences were found among the groups' listening and speaking anxiety mean scores on the posttests after controlling for the initial differences. Those experimental groups that received testing-the-limits and Lerntest approaches had a lower level of listening and speaking anxiety on the posttest. It was

---

[*] Corresponding author
[1] Department of English, Imam Khomeini International University, Qazvin, Iran; *a.zarei@hum.ikiu.ac.ir*
[2] Department of English, Imam Khomeini International University, Qazvin, Iran; *elham.shishegar@yahoo.com*

concluded that employing dynamic assessment models can decrease speaking and listening anxiety among EFL learners and enhance their productivity. The findings can have important implications for students, teachers and materials designers.

Keywords: anxiety, dynamic assessment, Learning Potential Measurement Approach, Lerntest Approach, Testing-the-Limits Approach

Introduction

Listening and speaking skills are usually neglected in Iranian public education, and teachers mostly spend their time on reading and writing skills. Due to learners' deficiency in listening and speaking skills, they usually feel anxious in classes involving oral activities (Zarei & Rezadoust, 2020). Anxiety is an affective factor that can hinder the learning process and demotivate learners. According to Horwitz (2001), there are three main types of anxiety, including state, trait, and situation-specific anxiety. Anxiety in a foreign language (FLA) is situation-specific anxiety (Aydin, 2008; Bailey et al., 1999). According to Abu-Rabia (2004), anxiety can have negative effects on all the four language skills. Although speaking is seen as the most anxiety-inducing skill, listening can also provoke high levels of anxiety, especially when it is incomprehensible (Young, 1992). Hence, second language teachers have always been seeking a way to help L2 learners deal with their listening and speaking anxiety.

One of the important variables that can potentially affect anxiety is assessment. Generally, classroom assessment can be classified into three categories: formative assessment (assessment for learning), summative assessment (assessment of learning), and self-assessment (assessment as learning) (Earl, 2006). Dynamic assessment (DA) is kind of assessment for learning. According to Hidri (2019), traditional assessments are used to assess the knowledge that a student has already acquired by experience while dynamic assessment determines the students' knowledge while they are being evaluated. Despite the fact that the concept of dynamic assessment has been around for more than 80 years (Poehner, 2008), it has not received much attention (Murphy & Maree, 2009), especially in Iranian language classes (Sanaeifard & Nafarzadeh Nafari, 2018).

There are two broad views regarding DA including interactionist versus interventionist (Lantolf & Poehner, 2011). Since there is a lack of standardization in

the interactionist approach and also the fact that it is a time-consuming, human-based assessment (Thouësny, 2010), three types of interventionist models, *Guthke's Lerntest Approach*, *Budoff's Learning Potential Measurement Approach*, and *Carlson and Wiedl's Testing-the-Limits Approach*, along with the conventional type of treatment were used in the present study to see if, and to what extent, they can decrease the level of L2 learners' listening and speaking anxiety. More specifically, this study attempted to address the following two research questions:

1. Do models of dynamic assessment differently affect L2 listening anxiety?
2. Do models of dynamic assessment differently affect L2 speaking anxiety?

Literature Review

Foreign Language Learning Anxiety

Generally, learning a language is regarded as a complex process, and many language learners do not feel comfortable learning or communicating in a foreign language. Many EFL learners report that they are stressed out in the process of learning a foreign language. In a study by Worde (1998), more than one third of language learners believed that they experienced high levels of anxiety.

The main sources of language anxiety, as stated by Horwitz et al. (1986), include concerns over negative evaluation and communication, and the effect of the test. Young (1991) enumerated several sources of anxiety in language learning. He also mentioned that the reasons behind language anxiety include interpersonal and personal anxiety, beliefs in learning a language, beliefs in language teaching, interactions based on instructor and learner relationship, procedures of classes, and target language test. Meanwhile, Chang and Read (2006) pointed out that three factors contribute to foreign language anxiety, including lack of confidence in oral language, prerequisite listening courses, and fear of test difficulty. Regardless of the sources of anxiety, a frequent observation in language classes is that anxiety is more vividly noticed in oral activities that include listening and speaking (Zarei & Rezadoust, 2020).

Listening Anxiety

Given the important role of listening, the issue of listening anxiety has been emphasized in the literature. Young (1992) pointed out that listening can cause great anxiety, especially among EFL learners. Many language learners have difficulties with this skill for different reasons such as invisible word boundaries and reduced forms (Ito, 2008). Besides, some language learners are not able to monitor the speech rate, or have limited processing skill (Buck, 2001). In addition, some EFL learners do not have a wide range of vocabulary, or they lack knowledge of the accent and topics in the target language (Buck, 2001). The effects of listening anxiety on the success of pedagogical activities has always been stressed by language scholars (Kimura, 2017). Some related studies have indicated that listening anxiety among foreign language learners influences their listening performance (Kimura, 2017). Young (1992) noted that when learners have difficulty figuring out the target language, they become anxious, especially when the teacher asks some questions about a listening file.

Furthermore, Vogely (1998) tried to identify the sources of foreign language listening anxiety in L2 pedagogical settings and found that it is related to the nature of input and the strategies and techniques that teachers use for teaching listening. Besides, Kim (2000) noted that the nature of listening input, like pronunciation, intonation, speed, the level of vocabulary, and the length of the text, plays a key role in listening anxiety. In the same vein, when students fail to distinguish the oral form of a sentence part, it has negative effects on their listening. Similarly, Yan (2005) showed that listening anxiety has a negative effect not only on listening comprehension but also on other language skills of students.

Speaking Anxiety

The relevant literature suggests that to many learners, speaking in a foreign language can generate a huge level of anxiety (Phillips, 1992). Ortega (2014) found that foreign language speaking anxiety is manifested in different ways, such as the high temperature of body, blood pressure, and the absence of eye contact. There are some sources of foreign language anxiety consisting of cultural, psychological, and linguistic factors. With respect to the low level of linguistic abilities, Ellis (2015) points out that learners with a lack of linguistic knowledge have indicated a higher

level of speaking anxiety.

Young (1992) listed six potential reasons for speaking anxiety including both interpersonal and personal factors like learners' beliefs on the language they intend to learn, the beliefs of teachers on the teaching process, interactions between teacher and learners, the procedures of the class and language tests. However, to date, those identified by Horwitz et al. (1986) have been regarded as the most influential. They listed three reasons for speaking anxiety including fear of negative evaluation, test anxiety, and communication apprehension.

Although it has been shown that speaking anxiety has a negative effect on L2 learners' performance (Horwitz, 2001; MacIntyre, 1995), there are only a few studies on how to control and reduce this anxiety (e.g., Safdari & Fathi, 2020). Therefore, there is a need for more investigations. It is important for language teachers to find ways to reduce the anxiety of learners in order to enhance their performance. One potential way is the use of dynamic assessment.

Dynamic Assessment (DA)

Developments in the methods of language teaching have provided parallel expansion in language assessment. Lantolf and Poehner (2011) noted that it is crucial to make a connection between education and assessment, and that these are combined well in a novel assessment approach called dynamic assessment.

Heywood and Lidz (2007) stated that DA mainly looks for distinguishing the unique characteristics of people, their weaknesses as well as their strengths, and their particular learning style to find effective methods for enhancing the effectiveness of learning for each individual. Actually, DA is an expansion-based process that shows the current abilities of learners, assists them to overcome any performance problems, and finds out their potential (Shrestha & Coffin, 2012).

The theory behind DA is the sociocultural theory (SCT). According to Murphy (2008), the founder of DA is Vygotsky, and the theory expands his common concept of the zone of proximal development (ZPD). According to this theory, learning can be defined as a process that happens in social contexts that are shaped by human intelligence. This process can affect the lives of learners and is focused on social and cultural domains. Since learning happens in a social context and is, therefore, socially mediated, interaction plays a key role in it. Vygotsky highlighted

the basic role of interaction in a social context in the cognitive expansion process (Kazemi et al., 2020). Given this theory, teaching is regarded as the process of assisting learners to expand their mental functions as well. Besides, according to this theory, teaching is seen as the process of working collaboratively with learners in such a way to develop their mental functions (Roebuck, 2001). In other words, assistance from more capable partners in interactions enable learners to improve their cognitive abilities and perform mental functions that they would not be able to perform without such assistance.

In DA, assessment and instruction come together and their combination occurs when intervention takes place within the evaluation process; its objective is to find out the abilities of learners and to guide them to higher levels of performance (Tabatabaee, et al., 2018). According to SCT, the fact that individuals respond to mediation that is geared to their ability level indicates that their cognitive functions have not completely expanded and that they need such assistance. Besides, suitable mediation permits individuals to control their performance independently, and this, in turn, leads to further expansion (Poehner, 2007).

The issue of mediation plays a critical role in the ZPD (Lantolf & Becket, 2009). Lantolf and Poehner (2011) pointed out that in DA, a particular type of mediated assistance is presented for learners. Thus, what causes an assessment to become static or dynamic is not the tool; it is whether the assessment of the course includes mediation or not (Poehner & Lantolf, 2005).

Generally, Lantolf and Poehner (2011) suggested two schools of thought concerning DA, referred to as interventionist and interactionist; both of them are regarded as a type of mediation. In interactionist DA, assistance will improve learning when the mediator and the learner are interacting and, thus, it is highly influenced by the learner's ZPD. However, interventionist DA is standardized and regarded as assistance according to the needs of learners to lead them to a pre-determined point quickly and effectively.

Sternberg and Grigorenko (2002) introduced two popular approaches of 'cake' and 'sandwich' for interventionist DA. The sandwich format mostly employs pretest, mediation, and posttest design to signify the improvement of learners and compare the pre and post-test. In the cake format, when learners fail to reply, the teacher intervenes throughout the test and employs preselected clues or prompts.

Thus, the learner is assessed during the process of intervention (Ahmadi Safa & Beheshti, 2018).

Furthermore, Poehner (2008) identifies interventionist models of DA as Lerntest Approach (Guthke's model), Learning Potential Measurement Approach (Budoff's model), Graduated Prompt Approach of Brown, and Testing-the-Limits Approach of Carlson and Wiedl. Because the interventionist and interactionist models of DA have already been compared in the literature in terms of their effectiveness on listening (Ahmadi Safa, & Beheshti, 2018) and speaking anxiety (Zarei & Rahmaty, 2021) and even other areas such as writing (Thouësny, 2010) and reading (Naeini, 2015), this study did not aim to compare the two general models. Instead, it focused on the different interventionist models, which are claimed to be standardized and lead learners to pre-specified objectives more effectively. For practicality reasons, from among the different interventionist models, the Learning Potential Measurement, the Lerntest, and the Testing-the-Limits models were employed.

Learning Potential Measurement. According to Poehner (2008), Budoff (1987) claimed that the possible effects of learners' background on their performance in tests can be stronger if learners are made familiar with the provided test and the strategies they need for problem solving in this process. In this respect, Budoff employed the sandwich format, a type of DA that borrowed most of its features from the classical 'pretest, treatment, post-test' research design. Such an approach made a crucial contribution to the claim of DA that cognitive abilities are dynamic since participants in Budoff's work reacted differently to mediation. Budoff classified learners based on their scores on the pretest and the posttest and showed that learners respond differently to instruction (Poehner, 2008).

Lerntest. Guthke and his associates extended the work of Budoff and designed a DA model referred to as the Lerntest (Poehner, 2008). Guthke (1982) stated that there is not one ZPD that is for general intelligence, but several ZPDs in different domains. His model moved the procedures of DA beyond the intelligence domain testing to consist of content areas like language. Guthke incorporated mediation into the test itself, which is in contrast to the work of Budoff. In the earlier forms of the model, learners were given only one kind of assistance when they gave a wrong answer; they were encouraged to try again. If learners still could

not give the right answer, the teacher provided it, so they moved on.

Later, Guthke developed five standardized hints. Pohner (2008) gives an example to show how LLT is conducted. According to Poehner, if a learner's initial response is not correct, the examiner provides some vague hints like "it's not correct, do you want to try it one more time". If s/he fails again, the examiner suggests a somewhat more explicit clue. And if the learner's attempt is still unsuccessful, the examiner presents a still more explicit clue. If it fails again, a rather explicit clue is offered. In the end, if this attempt also fails, the examiner presents the right pattern and elaborates why it is right. Then, the learner moves on to the next item of the text.

According to Guthke (1982), the list of standard points can be employed throughout different tests to run dynamic assessment. Such approaches employ tests in which dynamic assessment is the main focus; therefore, mediation between students and teachers is generally essential in such activities.

Testing-the-Limits. This approach is regarded as an alternative approach to the Lerntest; it was developed by Carlos and Wiedl (1978). These researchers also used standardized prompts and requested that learners explain how they came to their response. However, this approach is an extension of LLT. Carlson and Wiedl (1978) asked learners to elaborate on their reasoning regardless of whether or not their responses were correct so that they could find out the thought processes of learners completely. They believed that it is critical first to know how learners come to the answer and then decide if their answer is right (Zarei & Khojasteh, 2020).

Generally, testing-the-limits approach is mainly based on the belief that intra-individual differences in information processing result from personality and intellectual factors. This approach is mainly different from dynamic assessment approaches because particular mediations are combined within the process of testing (Safdari & Fathi, 2020). Measures of testing-the-limits approach seem to be restricted to general measures of the cognitive-based ability. This DA approach does not need modifications in the general structures or content of traditional tests. Instead, changes are at the heart of the testing process.

Related Studies

Several studies have been carried out on various aspects of the variables of this study. Alemi (2015) studied the effect of DA on L2 writing self-assessment and self-accuracy. She asked learners to write on seven topics. Data were collected through a pretest phase of DA when the students wrote a paragraph and evaluated their composition and filled out a questionnaire. In the final session, they wrote a paragraph again as the posttest and scored them according to the feedback they received. It was discovered that using DA led learners to a more accurate self-assessment. Moreover, engaging learners in assessment enhanced their autonomy.

In another study, Naeini (2015) compared the possible effects of Mediated Learning Experience and Graduated Prompt on L2 reading comprehension. The findings suggested that learners progressed gradually throughout the sessions of assessment, and DA assisted them to improve their reading comprehension.

Moreover, Abdolrezapour and Ghanbari (2021) examined the effect of employing self-regulated DA on enhancing self-regulation as well as listening comprehension. The members of the experimental group showed better scores in listening comprehension; their self-regulation also improved. In a similar study, Tabatabaee, et al. (2018) found that cumulative dynamic assessment had a significant effect on EFL learners' writing accuracy.

Ahmadi Safa and Beheshti (2018) investigated the effects of interventionist and interactionist DA on L2 listening comprehension. They found that interactionist dynamic assessment was more effective on developing listening comprehension. In the same vein, Safdari and Fathi (2020) examined the effect of DA on EFL learners' speaking fluency and accuracy. They concluded that DA had a significant positive effect on speaking accuracy, but it did not influence speaking fluency.

In their study, Kazemi and Tavassoli (2020) compared the effect of DA versus diagnostic testing on L2 speaking ability. They concluded that both dynamic and diagnostic assessments had a significant effect on EFL learners' speaking ability. Similar results were reported by Ahmadpour and Asadollahfam (2018). In addition, Wang (2015) reported the effectiveness of DA on listening skills.

Estaji (2019) investigated the short and long-term effect of DA approaches on the anxiety and oral performance of EFL learners. The findings suggested that both groups made progress in their speaking; they had better scores and lower

anxiety. Similarly, Zarei and Rezadoust (2020) investigated the effects of scaffolded and unscaffoled feedback on EFL learners' speaking self-efficacy and anxiety. They concluded that using both feedback types could decrease speaking anxiety and improve speaking self-efficacy in a foreign language. Meanwhile, Sohrabi and Ahmadi Safa (2020) reported that group dynamic assessment had a positive effect on learners' oral skill, and it reduced their anxiety.

Köroğlu (2019) examined the role of interventionist DA in speaking skills with a focus on ELT teachers. The findings showed that employing interventionist dynamic assessment positively influenced the participants speaking ability. In another study, Zarei and Khojaste (2020) compared the effect of three DA approaches including Group Dynamic Assessment, Intensive Mediated Learning Experience (MLE) and Learning Potential measurement on learning the lexical collocations of English. The findings indicated that MLE was more effective than both of the other models on the comprehension and production of lexical collocations.

Furthermore, Zarei and Rahmaty (2021) investigated the effects of interactionist versus interventionist DA models on L2 learners' willingness to communicate (WTC), foreign language anxiety (FLA) and perfectionism. They concluded that dynamic assessment reduces learners' FLA and reduces WTC and perfectionism. Besides, Woodrow (2006) found a negative relationship between L2 learners' speaking anxiety and speaking skill. The learners reported that interacting with native speakers was their main source of anxiety.

In their study, Capan and Karaca (2013) examined the relationship between education level, gender, and reading and listening anxiety. They concluded that the reason for these two skills (listening and reading) lies in the ambiguity in classroom objectives and that this ambiguity should be avoided in activities so that learners understand what they are supposed to do and for what reason.

Although there is no shortage of evidence on the effectiveness of DA models on improving language skills, few studies have checked their role in reducing anxiety. Even less research has been carried out on the comparativeness of effects of DA models on EFL learners' listening and speaking anxiety. Therefore, the goal of this study is to examine the effects of three kinds of DA on learners' L2 speaking and listening anxiety.

Method

Participants

The participants were 120 pre-intermediate Iranian English language learners (including members of both genders) with the age range of 18 to 35 selected through convenience sampling based on availability. They were in intact classes; therefore, Oxford Placement Test (OPT) was used ensure the homogeneity of the participants. They were all Persian speakers who had never lived or stayed longer than a 2-week period in any foreign country. They had the experience of studying English for almost six years.

Instrumentation

To accomplish the purpose of this study, the following instruments were used:

Oxford Placement Test (OPT). The OPT includes 60 items in a multiple-choice format. According to the scoring method and the criteria based on which to determine the level of proficiency of test takers, the students of this study were at pre-intermediate level, whose OPT score ranged between 30 and 39. The test has already been used extensively as a reliable instrument. Nevertheless, its reliability was re-estimated in the context of this study, and the index of internal consistency (Cronbach's alpha) turned out to be .90.

**Foreign Language Listening Anxiety Scale (FLLAS*).* This instrument, which was designed by Kim (2000), includes 33 items; each item is on a five-point Likert type scale that ranges from "Strongly Disagree" (1) to "Strongly Agree" (5). The possible range of scores is from 33 to 165. High scores on this questionnaire indicate high levels of listening anxiety. Cronbach's alpha was used to estimate the internal consistency reliability index of this questionnaire in the new context of the present study, and it turned out to be .88.

**Foreign Language Classroom Speaking Scale (FLCSS*).* Horwitz et al's.*,* (1986) (FLCSS) consists of 12 items on a 5-point Likert type scale ranging from "Strongly Disagree" (1) to "Strongly Agree" (5). The scores could possibly range from 12 to 60. The higher the score, the higher the level of speaking anxiety. The internal consistency of this scale in the context of this study was assessed to be .85.

Procedure

The participants were selected from among the pre-intermediate students of Kish-e-mehr language institute in Qazvin. They selection was done through convenience sampling that was largely on the basis of availability. Four intact classes of learners were selected, each with 30 students.

Both questionnaires were administered to the learners in the first session. Then, treatment was given for ten sessions. Three groups were treated by three models of dynamic assessment as experimental groups. The first group was randomly selected to be treated with Budoff's LPM Approach which included pretest, treatment, and posttest. Learners were taught strategies of problem-solving through a standardized procedure. In fact, learners were grouped on the basis of the differences that they showed in their scores on the pretest and posttest. There were three categories: 'high scorers', who performed well on the pretest; 'gainers', who performed well after receiving the treatment; and 'nongainers', who did not perform well on the either the pretest or the posttest. The second group was treated with Guthke's Lerntest Approach. In applying this approach, there were five standardized hints. When a leaner's initial attempt was not successful, s/he was given a vague clue. When the second attempt was also unsuccessful, s/he was provided with a more explicit hint. When the response was still incorrect, the examiner offered a still more explicit clue. When s/he still failed, a very explicit hint was offered, and finally, when the last attempt failed again, the tester provided the correct answer and explained the reason for the answer. The third group was treated with Testing-the-Limits Approach. Based on the principles of this approach, the examiner employed standardized hints (like in the previous treatment group) again and also asked the learners to explain the reason for their responses. The fourth group acted as the control group. In this group, the participants were not treated with any dynamic assessment model; instead, they were taught conventionally. Like the other three treatment groups, this process continued for ten sessions.

In the last session of the treatment period, in the twelfth session, the same listening and speaking anxiety questionnaires were administered again to all the learners in the four groups. Lastly, the scores of all the four groups on both the pre- and the post-test of listening and speaking anxiety were compared and analyzed to address the two research questions of the study.

Data Analysis

To answer the first research question probing the comparative effects of three models of dynamic assessment on L2 listening anxiety, one-way analysis of covariance (one-way ANCOVA) was used. The scores of the students on the pre-test were taken as the covariate in order to 'control' for the already existing differences among the groups. Another one-way ANCOVA was used to address the second question exploring the effect of different models of dynamic assessment on L2 speaking anxiety.

Results and Discussion

Research Question 1

The purpose of the first research question was to find out whether or not there were any meaningful differences among the models of dynamic assessment with regard to their effectiveness on listening anxiety after controlling for the initial differences. An ANCOVA was used to address this question. Before so doing, its assumptions were checked. The results of checking the assumption of the linearity of the relationship between the posttest of listening anxiety (the dependent variable) and the pretest of listening anxiety (the covariate) are displayed in Figure 1. As shown in the scatter plot, there are four straight lines in the scatterplot of the dependent variable as well as the covaiate scores of the four groups of the study. These straight lines indicate that the linearity assumption was not violated.

Figure 1

Scatter plot for listening anxiety scores
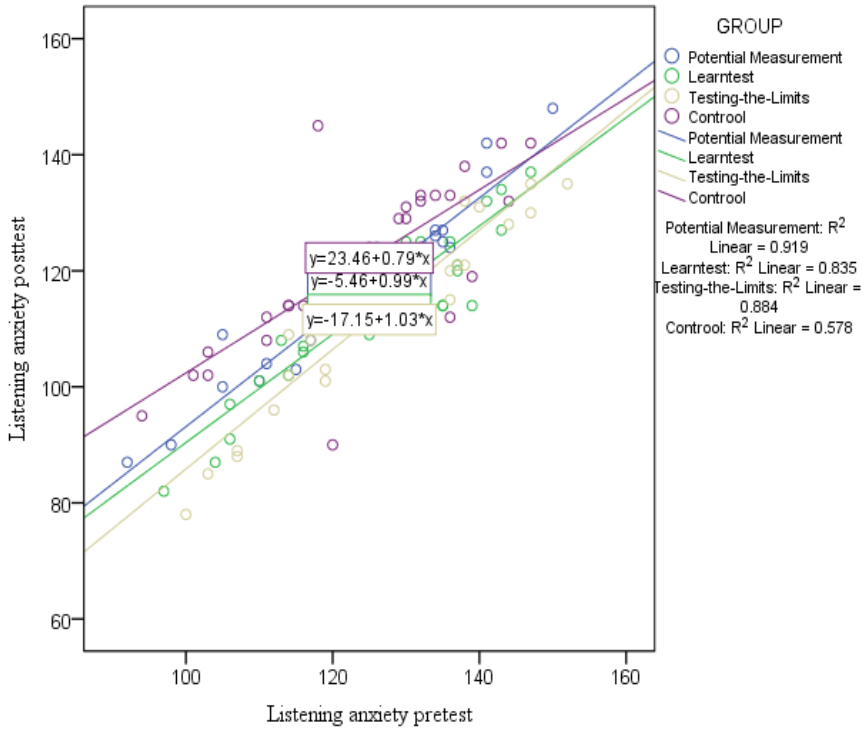
As it is evident from Table1, the significance level of the Levene's test (.19) was greater than (.05). This means that we can safely assume that the variances of scores in the listening anxiety questionnaire across the groups enjoys homogeneity.

Table 1

Levene's test of equality of error variances for listening anxiety

Levene Statistic	df1	df2	Sig.
1.596	3	116	.194

Table 2 indicates that the interaction between the pretest of listening anxiety and the treatment ($F_{(3, 112)}$ = 1.46, p = .23) did not reach statistically significance level. This shows that the assumption of homogeneity of regression slopes was also met.

Table 2

Homogeneity test of regression slopes for listening anxiety

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	19481.19	7	2783.02	69.81	.00	.81
Intercept	.36	1	.36	.009	.92	.00
Group * Pretest	174.99	3	58.33	1.46	.22	.03
Error	4464.77	112	39.86			
Total	1643774.00	120				
Corrected Total	23945.96	119				

Table 3 summarizes the descriptive statistics for the listening anxiety scores. The mean score of listening anxiety in the potential measurement group ($M = 123.77$, $SD = 13.23$), Lerntest group ($M = 125.40$, $SD = 12.91$), testing-the-limits ($M = 126.33$, $SD = 13.64$), and control group ($M = 123.70$, $SD = 13.56$) are not far from one another on the pretest; however, when the posttest results are compared, the average of the testing-the-limits group ($M = 112.97$, $SD = 14.94$) is the lowest of all, followed by the Lerntest group ($M = 114.10$, $SD = 13.37$), the potential measurement group ($M = 116.57$, $SD = 13.61$), and the control group ($M = 121.10$, $SD = 14.08$).

Table 3

Descriptive statistics of listening anxiety scores

Group	N	Pre-test Mean	Post-test Mean
Potential measurement	30	123.77	116.57
Lerntest	30	125.40	114.10
Testing-the-limits	30	126.33	112.97
Control	30	123.70	121.10

Table 4 contains the result of ANCOVA. After considering the scores of the participants on the pretest of listening anxiety, the differences among the study groups were significant with respect to their scores on the listening anxiety posttest,

($F_{(3, 115)}$ = 16.20, p = .000, partial eta squared = .28). Besides, a strong relationship ($F_{(1, 115)}$ = 449.51, $p < .005$) was detected between the pre- and posttest scores of listening anxiety. This shows that the scores on listening anxiety on the pretest influenced the listening anxiety scores obtained on the posttest, with an effect size of .79.

Table 4

ANCOVA results for listening anxiety

Source	Type III SS	df	MS	F	Sig.	Partial η^2
Corrected Model	19306.200a	4	4826.550	119.630	.000	.806
Intercept	.238	1	.238	.006	.939	.000
Pretest	18135.966	1	18135.966	449.513	.000	.795
Group	1960.968	3	653.656	16.201	.000	.285
Error	4639.767	115	40.346			
Total	1643774.000	120				
Corrected Total	23945.967	119				

To further explore the possible differences among the average scores of the four groups on listening anxiety, pairwise comparisons were made (Table 5). Pairwise comparisons revealed a statistically meaningful difference in the listening anxiety mean scores between the control group and the three experimental groups. Moreover, a statistical difference was observed between the Lerntest and potential measurement groups, with the mean difference of 3.99 in favor of the potential measurement group. Likewise, pairwise comparisons yielded a significant difference with regard to the listening anxiety between the testing-the-limits and potential measurement groups, with the mean difference of almost 6.00 in favor of the potential measurement group. However, the difference between the testing-the-limits and the Lerntest groups was statistically insignificant.

Table 5

Pairwise comparisons for listening anxiety

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Sig.
Lerntest	Potential measurement	-3.993*	.015
Testing-the-limits	Potential measurement	-5.998*	.000
Testing-the-limits	Lerntest	-2.005	.224
Control	Potential measurement	4.596*	.006
Control	Lerntest	8.588*	.000
Control	Testing-the-limits	10.594*	.000

Research Question 2

The second question of this study was intended to find out whether or not there were any meaningful differences among the three models of DA with respect to their effectiveness on speaking anxiety. Another ANCOVA was utilized to answer this research question. To check whether or not there was a linear relationship between the speaking anxiety pretest and posttest, a scatter plot was drawn, which showed that the linearity assumption was met. The probability level of the Levene's test (Table 6) confirmed that the speaking anxiety scores enjoyed homogeneity in their variance.

Table 6

Levene's test results for speaking anxiety

Levene Statistic	df1	df2	Sig.
.115	3	116	.874

According to Table 7, the interaction between the grouping variable and the pretest of speaking anxiety (Group * Pretest) did not turn out to be statistically significant ($F_{(3, 112)} = 2.10$, $p > .05$) showing that there was no serious threat to the assumption of homogeneity of regression slopes.

Table 7

Homogeneity test of regression slopes for speaking anxiety

Source	Type III SS	df	MS	F	Sig.	Partial η^2
Corrected Model	11030.476	7	1575.782	550.637	.000	.972
Intercept	1.675	1	1.675	.585	.446	.005
Group * Pretest	18.015	3	6.005	2.098	.104	.053
Error	320.516	112	2.862			
Total	205839.000	120				
Corrected Total	11350.992	119				

Having made sure that every assumption was met, the researchers went on to use the one-way ANCOVA. The speaking anxiety scores of the four groups are summarized in Table 8. It shows that the pretest average scores of speaking anxiety in the potential measurement group ($M = 43.17$, $SD = 11.39$), Lerntest group ($M = 43.93$, $SD = 9.94$), testing-the-limits ($M = 44.03$, $SD = 9.28$), and control group ($M = 42.43$, $SD = 10.95$) are rather close to one another; however, when it comes to the posttest, the average of the testing-the-limits group ($M = 39.73$, $SD = 9.33$) is the lowest of all, followed by the Lerntest group ($M = 39.87$, $SD = 9.67$), the potential measurement group ($M = 40.40$, $SD = 10.44$), with the control group coming last ($M = 41.03$, $SD = 10.03$).

Table 8

Descriptive statistics of speaking anxiety scores

Group	N	Pre-test Mean	Post-test Mean
Potential measurement	30	43.17	40.40
Lerntest	30	43.93	39.87
Testing-the-limits	30	44.03	39.73
Control	30	42.43	41.03

Table 9 summarizes the main ANCOVA results. It shows that, after controlling for the initial differences in speaking anxiety scores, there was still a

meaningful difference among the four groups' speaking anxiety mean scores on their posttest, $(F_{(3, 115)} = 16.84, p < .005$, partial eta squared = .30). In addition, the pre-intervention scores were a significant covariate of speaking anxiety scores on the posttest $(F_{(1, 115)} = 3730.27, p < .005)$.

Table 9

ANCOVA results for speaking anxiety

Source	Type III SS	df	MS	F	Sig.	Partial η^2
Corrected Model	11012.461	4	2753.115	935.243	.000	.930
Intercept	.354	1	.354	.120	.729	.001
Pretest	10980.970	1	10980.970	3730.275	.000	.930
Group	148.756	3	49.585	16.844	.000	.305
Error	338.530	115	2.944			
Total	205839.000	120				
Corrected Total	11350.992	119				

Pairwise comparisons (Table 10) yielded a meaningful differences in the speaking anxiety meanscores between the control group and the three treatment groups. Furthermore, a significant difference could be seen between the Lerntest and potential measurement groups, in favor of the potential measurement group. Moreover, pairwise comparisons revealed a difference that reached statistical significance between the testing-the-limits and potential measurement groups in favor of the potential measurement group. However, there was only a negligible and statistically insignificant difference between the testing-the-limits and Lerntest groups.

Table 10

Pairwise comparisons for speaking anxiety

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.
Lerntest	Potential measurement	-1.255*	.437	.005
Testing-the-limits	Potential measurement	-1.505*	.437	.001
Testing-the-limits	Lerntest	-.251	.438	.568
Control	Potential measurement	1.298*	.438	.004
Control	Lerntest	2.554*	.438	.000
Control	Testing-the-limits	2.805*	.438	.000

Discussion

This study showed that the selected models of DA were differentially effective on L2 speaking anxiety. In fact, it was observed that the testing-the-limits model was the most effective model of all. These findings support those of Sohrabi and Ahmadi Safa (2020) in which they concluded that employing group dynamic assessment can enhance the oral productivity of EFL learners as well as reducing their speaking anxiety.

Also, the findings are compatible with those of Ahmadpour and Asadollahfam (2018) who found the effectiveness of DA on the oral performance of EFL learners. It can be argued that better oral performance may result in a lower level of speaking anxiety. Furthermore, this finding of the study confirm that of Estaji (2019) and Kazemi and Tavassoli (2020) who found that employing both immediate and delayed dynamic assessment approaches can reduce the speaking anxiety of EFL learners and improve their oral performance.

This study also found meaningful differences among the selected models of DA with regard to their effect on L2 listening anxiety. Such a finding is congruent with that of Abdolrezapour and Ghanbari (2021), Ahmadi Safa and Beheshti (2018), and Wang (2015) who concluded that employing self-regulated dynamic assessment can reinforce EFL learners' listening comprehension. Since better performance in one skill is normally associated with a lower level of anxiety in that skill (Sohrabi & Ahmadi Safa, 2020), it can be said that the results of their study were compatible with the findings of the current study. Besides, these findings correlate with those of

Zarei and Rahmaty (2021), who showed that both interactionist and interventionist dynamic assessment reduced foreign language anxiety of EFL learners.

Meanwhile, the finding that both the Lerntest and the Testing-the-limits models were more effective than the potential measurement model on reducing both speaking and listening anxiety may be accounted for by the fact that in the potential measurement model, which follows the sandwich format, there is only one stage of feedback, whereas in both of the former models, there are a series of staged feedback; specifically five stages of feedback in more recent versions of the models.

Another reason that could potentially explain the superiority of the Lerntest and the Testing-the-limits models over the potential measurement model in reducing speaking and listening anxiety is that in the potential measurement model of DA, mediation is almost completely focused on the learners' cognitive characteristics, assuming that learners with different cognitive characteristics respond differently to mediation. However, in the Lerntest model, and the Testing-the-limits model which is an extension of the Lerntest, the assumption is that learners do not have just a single cognitive ZPD; rather, they have a number of domain-specific ZPDs that need to be considered when giving mediation. In other words, in these models, learners are expected to be involved in learning in multidimensional ways. This means that in addition to the cognitive characteristics of learners, their emotional and social characteristics are also considered in the mediation phase. Since both listening and speaking anxiety are affective variables, it seems that theoretically the results of this study make sense.

One final reason accounting for the effectiveness of the Lerntest and the Testing-the-limits models in comparison to the potential measurement model in reducing speaking and listening anxiety may be that both of the mentioned models are more sensitive to individual differences among learners in responding to mediation. In other words, they offer more individualized and personalized mediation to learners. As a result, learners can benefit from the mediation more.

Conclusion

From the results of the present study, it is concluded that although all the three models of dynamic assessment can have more positive effect than conventional testing and instruction on reducing speaking and listening anxiety,

each of these models appears to be more useful for specific purposes. Generally, employing such models in EFL pedagogical settings can reduce speaking and listening anxiety among EFL learners and enhance their productivity.

The potential measurement model focuses on the cognitive abilities of learners and makes a critical contribution to the claim of DA that cognitive abilities are dynamic. Lerntest assumes that each person can have multiple ZPDs in different domains, not just one ZPD for general intelligence. In Testing-the-limits, the teacher asks learners to explain how they got to their answer irrespective of whether their responses are correct or incorrect so that they can find out the thought processes of learners thoroughly. This model can present more information given leaners' abilities. It is employed to evaluate the limits of learners' abilities by integrating different procedures that causes higher levels of performance.

In sum, the focus of dynamic assessment is on the process of learning not its final product. That is why employing this assessment was successful in reducing the anxiety of students. According to the sociocultural theory, human learning can be seen as a type of mediated learning. It means that employing mediation assists the cognitive development of humans independent from other regulations. In other processes of regulation, there is no graduated prompt to move learners toward effective processing. Besides, in models and approaches in which teachers emphasize the final outcome of learners, students have high levels of stress and anxiety since they try to do their best to score well. The use of relevant DA models can be both effective in preventing the occurrence of such anxiety and helpful in reducing it when anxiety has already been generated.

Therefore, the results of this study can extend the literature on the effectiveness of DA models in reducing listening and speaking anxiety of EFL learners. The effectiveness of DA models can be explained with the assumption that while students at varying levels of language proficiency complete the same set of activities, each learner actually needs particular, and somehow individualized, feedback adjusted to their personal needs. As a result, it can be deduced that this feature of DA models stresses learners' needs. Thus, employing DA can be seen as a more ethical and equitable way of assessing learners (Shabani, 2018).

The findings of this study can also inform language teachers about how to treat learners to provide the best learning conditions by bringing the level of

listening and speaking anxiety to a minimum. It is undeniable that oral activities in foreign language learning classes are anxiety-inducing. The knowledge of which model of DA can reduce listening and speaking anxiety can hugely facilitate teachers' job in listening and speaking classes. Moreover, materials developers can include materials in course books that encourage and guide both teachers and learners to engage in the right kind of activities, thus reducing the anxiety of the learners. Nevertheless, the limitations of this study including the number of participants and the duration of treatments, call for more studies in this area and a greater level of care to be exercised in generalizing these findings.

References

Abdolrezapour, P. & Ghanbari, S. (2021). Enhancing learning potential score in EFL listening comprehension and self-regulation through self-regulated dynamic assessment procedures. *Language Testing in Asia, 2*(10), 22-34. https://doi.org/10.1186/s40468-021-00126-5

Abu-Rabia, S. (2004). Teachers' role, learners' gender differences, and FL anxiety among seventh-grade students studying English as a FL, *Educational Psychology, 24*(5), 711-720. https://psycnet.apa.org/doi/10.1080/0144341042000263006

Ahmadi Safa, M. & Beheshti, S. (2018). Interactionist and interventionist group dynamic assessment (GDA) and EFL learners' listening comprehension development. *Iranian Journal of Language Teaching Research, 6*(3), 37-56. https://doi.org/10.30466/ijltr.2018.120600

Ahmadpour, L. & Asadollahfam, H. (2018). The effects of dynamic assessment and gender on EFL learners' oral narrative task performance. *International Journal of Language Testing and Assessment, 1*(1), 12-24. https://doi.org/10.24815/ijolta.v1i1.9773

Alemi, M. (2015). The Impact of dynamic assessment on Iranian EFL students' writing self-assessment. *TELL, 9*(1), 145-169. https://doi.org/10.22132/tel.2015.53735

Aydin, S. (2008). An investigation on the language anxiety and fear of negative evaluation among Turkish EFL learners, *Asian EFL Journal*, *30*(1), 421-444.

Bailey, P., & Daley, C. E. (1999). Foreign language anxiety and learning style. *Foreign Language Annals*, *32*(1), 63-76.

Bailey, P., Onwuegbuzie, A. J., & Daley, C. E. (1999). Anxiety about foreign language students in French, Spanish, and German classes. *Psychological Reports*, *8*(2), 1007-1010.

Bekka, K.G. (2010). Dynamic assessment for learning potential: A shift in the focus and practice of evaluating Japanese oral proficiency. *Japanese Journal of Education, 2(1)*, 53-66.

Buck, G. (2001). *Assessing listening*. Cambridge University Press.

Budoff, M. (1987). Learning potential as a supplementary testing procedure. In J. Hellmuth (ed.) *Learning disorders,* (Vol. 3). Special Child.

Capan, S. A., & Karaca, M. (2013). A comparative study of listening anxiety and reading anxiety. *Procedia-Social and Behavioral Sciences, 70*(5), 1360-1373. http://dx.doi.org/10.1016/j.sbspro.2013.01.198

Carlson, J. S., & Wiedl, K. H. (1978). Use of testing-the-limits procedures in the assessment of intellectual capabilities in children with learning difficulties. *American Journal of Mental Deficiency, 82*(6), 559-564.

Chang, A., C. S., & Read, J. (2006). The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly, 40*(5), 375–394. https://doi.org/10.2307/40264527

Earl, L. (2006). *Rethinking classroom assessment with purpose in mind.* Western Northern Canadian Protocol.

Ellis, R. (2015). *Understanding second language acquisition* (2nd Ed.). Oxford University Press.

Estaji, M. (2019). The Immediate and delayed effect of dynamic assessment approaches on EFL learners' oral narrative performance and anxiety. *Educational Assessment, 24*(2), 135-154. https://doi.org/10.1080/10627197.2019.1578169

Fatemipour, H., & Jafari, F. (2015). The Effect of dynamic-assessment on the development of passive vocabulary of intermediate EFL learners. *Journal of Educational and Management Studies, 5*(1), 41-51.

Guthke, J. (1982). The learning test concept and its application in practice. In C. S. Lidz, & J. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications* (pp. 17-69). Elsevier.

Haywood H. C. & Lidz C. S. (2007). *Dynamic assessment in practice: Clinical and educational applications.* Cambridge University Press.

Hidri, S. (2019). Static vs. dynamic assessment of students' writing exams: a comparison of two assessment modes. *International Multilingual Research Journal*, *13*(4), 239-256. http://doi.org/10.1080/19313152.2019.1606875

Horwitz, E. K. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics*, *21*, 112-126. https://doi.org/10.1017/S0267190501000071

Horwitz, M. B., Horwitz, E. K. & Cope, J. A. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125-132. https://doi.org/10.2307/327317

Ito, N. (2008). *Exploring the nature of language anxiety: Experiences of nonNative English speaking college students in the United States.* University of New Orleans, Theses and Dissertations. 821. https://scholarworks.uno.edu/td/821

Kazemi, N. & Tavassoli, K. (2020). The Comparative Effect of dynamic vs. diagnostic assessment on EFL learners' speaking ability. *Research in English Language Pedagogy, 3*(4), 45-59. https://doi.org/10.30486/relp.2019.1878561.1155

Kazemi, A., Bagheri, M., & Rassaei, E. (2020). Dynamic assessment in English classrooms: Fostering learners' reading comprehension and motivation. *Cogent Psychology*, *7*(1), 1-10. http://doi.org/10.1080/23311908.2020.1788912

Kim, J. H. (2000). *Foreign language listening anxiety: A study of Korean students learning English.* Unpublished doctoral dissertation. University of Texas.

Kimura, H. (2017). Foreign language listening anxiety: Its dimensionality and group differences. *JALT, 30*(2), 173-196. http://dx.doi.org/10.37546/JALTJJ30.2-2

Köroğlu, Z. C. (2019). Interventionist dynamic assessment's effects on speaking skills testing: case of ELT teacher candidates. *Advances in Language and Literary Studies, 10*(3), 23-31. https://doi.org/10.7575/aiac.alls.v.10n.3p.23

Lantolf, J. P., & Poehner, M. E. (2011). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research, 15*(1), 11-33. http://dx.doi.org/10.1177/1362168810383328

Lantolf, J., & Beckett, T. (2009). Sociocultural theory and second language acquisition. *Language Teaching, 42*(4), 459-475. http://dx.doi.org/10.1017/S0261444809990048

Lee, I. (2007). Assessment for learning: Integrating assessment, teaching, and learning in the ESL/EFL writing classroom. *Canadian Modern Language Review, 64*(1), 199-213. http://dx.doi.org/10.3138/cmlr.64.1.199

MacIntyre, P. D. (1995). How does anxiety affect second language learning? A reply to Sparks and Ganschow. *The Modem Language Journal*, *79*(1), 90-99. https://doi.org/10.2307/329395

Murphy, R. (2008). Dynamic assessment precursors: Soviet ideology and Vygotsky. *The Irish Journal of Psychology, 29*(4), 193-233. http://dx.doi.org/10.1080/03033910.2008.10446285

Murphy, R. & Maree, D. J. F. (2009). Revisiting core issues in dynamic assessment. *South African Journal of Psychology*, *39*(4), 420-431. http://dx.doi.org/10.1177/008124630903900404

Naeini, J. (2015). A Comparative study of the effects of two approaches of dynamic assessment on the reading comprehension of Iranian EFL learners. *International Journal of Applied Linguistics and English Literature, 3*(10), 33-43. http://dx.doi.org/10.7575/aiac.ijalel.v.4n.2p.54

Ortega, L. (2014). *Understanding second language acquisition*. Routledge.

Pahlavani, P., & Maftoon, P. (2015). The impact of using computer-aided argument mapping (CAAM) on the improvement of Iranian EFL learners' writing self-regulation. *The Journal of Teaching Language Skills (JTLS)*, *7*(2), 127-156.

Phillips, E. M. (1992). The effects of language anxiety on students oral test performance and attitudes. *The Modern Language Journal, 76*(1), 14-26. https://doi.org/10.1111/j.1540-4781.1992.tb02573.x

Poehner, M. E, & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research, 9*(3), 233–265.

https://doi.org/10.1191/1362168805lr166oa

Poehner, M. E. (2007). Beyond the test: L2 dynamic assessment and the transcendence of mediated learning. *The Modern Language Journal, 91*(3), 323–340. https://psycnet.apa.org/doi/10.1111/j.1540-4781.2007.00583.x

Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. Springer

Roebuck, R. F. (2001). Teaching composition in the college level foreign language class: Insights and activities from sociocultural theory. *Foreign Language Annals, 34*(3), 206-215.

Safdari, M. & Fathi, J. (2020). Investigating the role of dynamic assessment on speaking accuracy and fluency of pre-intermediate EFL learners. *Cogent Education, 2*(3), 23-45. https://doi.org/10.1080/2331186X.2020.1818924

Sanaeifard, S. H. & Nafarzadeh Nafari, F. (2018). The effects of formative and dynamic assessments of reading comprehensions on intermediate EFL learners' test anxiety. *Theory and Practice in Language Studies*, *8*(5), 533-540. http://dx.doi.org/10.17507/tpls.0805.12

Shabani, K. (2018). Group dynamic assessment of L2 learners' writing abilities. *Iranian Journal of Language Teaching Research*, *6*(1), 129-149. https://doi.org/10.30466/ijltr.2018.20494

Shrestha, P., & Coffin, C. (2012). Dynamic assessment, tutor mediation and academic writing development. *Assessing Writing, 17*(1), 55-70. http://dx.doi.org/10.1016/j.asw.2011.11.003

Sohrabi, S., & Ahmadi Safa, M. (2020). Group dynamic assessment and EFL learners' oral production, motivation, and classroom anxiety. *English Teaching & Learning, 44(*2), 1-27. http://doi.org/10.1007/s42321-020-00054-2

Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential.* Cambridge University Press.

Tabatabaee, M., Alidoust, A., & Sarkeshikian, A. H. (2018). The effect of interventionist and cumulative group dynamic assessments on EFL learners' writing accuracy. *Applied Linguistics Research Journal, 2*(1), 1-13. http://dx.doi.org/10.14744/alrj.2018.36854

Thouësny, S. (2010). Assessing second language learners ' written texts: An interventionist and interactionist approaches to dynamic assessment. *Educational Studies, 4*(3), 3517-3522.

Veresov, N. N. (2017). ZBR and ZPD: Is there a difference? *Cultural-Historical Psychology, 13*(1), 23-36. http://doi.org/10.17759/chp.2017130102

Vogely, A. J. (1998). Listening comprehension anxiety: Students' reported sources and solutions. *Foreign Language Annals,* *31*(1), 67–80. http://dx.doi.org/10.1111/j.1944-9720.1998.tb01333.x

Vygotsky, L. (1978). *Mind in society*. Harvard University Press.

Young, D. J. (1991). An investigation of students' perspective on anxiety and speaking. *Foreign Language Annals, 23*(6), 539-553. http://dx.doi.org/10.1111/j.1944-9720.1990.tb00424.x

Young, D. J. (1992). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal, 75*(2), 426-439.

Wang, W (2015). Teaching English as an international language in China: investigating university teachers' and students' attitudes towards China English*. System, 53*, 60–72. https://doi.org/10.1016/j.system.2015.06.008

Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC Journal, 37*(3), 308-328. *https://doi.org/10.1177/0033688206071315*

Worde, R. V. (1998). Students' perspectives on foreign language anxiety. *Language Learning, 44*, 283-305.

Zarei, A., & Khojasteh, A. (2020). Models of dynamic assessment affecting the learning of English lexical collocations. *Journal of Language Horizons*, *4*(2), 239-259. https://doi.org/10.22051/lghor.2020.29463.1229

Zarei, A., & Rahmaty, H. (2021). The effects of interactionist and interventionist dynamic assessment on EFL students' perfectionism, willingness to communicate, and foreign language anxiety. *International Journal of Language Testing*, *11*(2), 13-33.

Zarei, A. A. & Rezadoust, H. (2020). The effects of scaffolded and unscaffolded feedback on speaking anxiety and self-efficacy. *Journal of Modern Research in English Language Studies, 7*(4), 111-132. http://dx.doi.org/10.30479/jmrels.2020.13464.1655