

**Effect of Interactionist, Group and Computerized Dynamic Assessment  
on Iranian EFL Learners' Listening Comprehension**

**Mahshid Ghenaat**

English Department, Shahrekord Branch, Islamic Azad University, Shahrekord, Iran;  
[2ghenaat@gmail.com](mailto:2ghenaat@gmail.com)

**Fariba Rahimi Esfahani**

Corresponding author: English Department, Shahrekord Branch, Islamic Azad University,  
Shahrekord, Iran  
[rahimi\\_fariba@yahoo.com](mailto:rahimi_fariba@yahoo.com)

**Sajad Shafiee**

English Department, Shahrekord Branch, Islamic Azad University, Shahrekord;  
[shafiee\\_sajad@yahoo.com](mailto:shafiee_sajad@yahoo.com)

**Mehrdad Sepehri**

English Department, Shahrekord Branch, Islamic Azad University, Shahrekord  
[m.sepehri@iaushk.ac.ir](mailto:m.sepehri@iaushk.ac.ir)

**ABSTRACT**

This experimental study aimed to make comparisons among the three models of dynamic assessment; namely, Interactionist Dynamic Assessment (I-DA), Group Dynamic Assessment (G-DA), and Computerized Dynamic Assessment (C-DA), in order to see if there are any differences among them with respect to their effects on the listening comprehension ability of female Iranian intermediate EFL learners. The obtained results revealed that, first of all, the three models of assessment affected the participants' listening comprehension and, secondly, Group Dynamic Assessment (G-DA) had a more positive impact on their listening comprehension ability in comparison with the other two models. The findings may suggest an alteration in the traditional models of listening comprehension assessment which accentuates no opportunities for learner-teacher interaction.

*Keywords:* Interactionist Dynamic Assessment (I-DA) - Group Dynamic Assessment (G-DA) - Computerized Dynamic Assessment (C-DA) - listening comprehension - dynamic assessment

**RESUMEN**

La investigación experimental sobre la que informa este artículo buscó comparar tres modelos de evaluación dinámica, la Evaluación Dinámica Interactiva, la Evaluación Dinámica Grupal y la Evaluación Dinámica Computarizada, para detectar diferencias entre ellas con respecto a sus efectos sobre la habilidad de comprensión auditiva en estudiantes de inglés como lengua extranjera. Los sujetos fueron mujeres iraníes con un conocimiento intermedio del idioma. Los resultados obtenidos demuestran que, primero, estos tres modelos de evaluación afectaron la comprensión auditiva y, segundo, la Evaluación Dinámica Grupal tuvo un impacto positivo mayor en la habilidad de comprensión auditiva. Estos resultados pueden sugerir una alteración en los modelos tradicionales de evaluación de comprensión auditiva, modelos que no suelen dar oportunidades para la interacción docente-alumno.

*Palabras claves:* Evaluación Dinámica Interactiva – Evaluación Dinámica Grupal – Evaluación Dinámica Computarizada – comprensión auditiva – evaluación dinámica

### Introduction

Traditional methods of assessment have fallen short in providing an inclusive view of the learners' performance in terms of developing language skills. Many testing experts (e.g. (Alderson, Percsich, and Szabo, 2000) have long welcomed standardized diagnosis tests as valid, reliable, and practical despite their limited scope in uncovering the latent cognitive and metacognitive skills and sub-skills that dynamic assessment (DA) is able to uncover. Alternatively, linking learning to a socio-cultural context is another approach that calls for participants to be part of a learning activity where much support is needed to diagnose the language proficiency of these learners. Basically, according to Vygotsky (1986), learners can show signs of successful learning in the presence of a more competent peer who can engage them in mediation activities and tasks using specific strategies and skills to move from their actual to the proximal zone of development (as cited in Lantolf and Poehner, 2007).

### Types of Dynamic Assessment

#### *Interactionist Dynamic Assessment (I-DA)*

The interactionist design, or “teaching within test format”, provides learners with mediation in the shape of hints, prompts and instruction; it is generally sequenced from implicit to the most explicit (Shabani, 2012) or from general to specific. “Interactionist dynamic assessment is based on Vygotsky’s focus on cooperative dialoging. In this view, the interaction is between the teacher and student and the teacher helps students whenever it is necessary” (Grigorenko, 2009, p. 38).

#### *Group Dynamic Assessment (G-DA)*

G-DA is referred to as the group activity which happens in the sociocultural context where people or students interact with one another. Poehner (2009) defined G-DA based on Vygotsky’s approach where the optimal space for improving learning is the one between what each learner can individually do and what they can perceive with the help of others, especially those who are more capable, such as teachers.

#### *Computerized Dynamic Assessment (C-DA)*

Computerized Dynamic Assessment (C-DA) is an ongoing version of DA, which provides students with automatic mediations through computers. Computerized dynamic assessment (C-DA) includes mediation in the assessment process. Proponents of dynamic assessment (DA) in general and C-DA in particular argue that the goals of DA are in congruence with the concept of validity that underscores the social consequences of test use and the integration of learning and assessment (Sternberg & Grigorenko, 2002; Poehner, 2008; Alonazi, 2017)

Simultaneous evaluation of the impact of different types of dynamic assessment on EFL learners’ listening comprehension ability has never been conducted as far as the related literature discloses. Thus, the aim of the present study was to make comparisons among three models of DA; namely I-DA, G-DA, and C-DA, in order to see if there are any differences among them with respect to their effects on the listening comprehension of EFL learners. Consequently, the following research questions were addressed:

- RQ1.** Does using I-DA have any significant effect on Iranian EFL learners' listening comprehension?
- RQ2.** Does using G-DA have any significant effect on Iranian EFL learners' listening comprehension?
- RQ3.** Does using C-DA have any significant effect on Iranian EFL learners' listening comprehension?
- RQ4.** Is there any significant difference among the effects of interactionist, group, and computerized dynamic assessment on Iranian EFL learners' listening comprehension?

## Method

### *Design of the Study*

To achieve the purposes of the study, a quantitative approach was employed, with a quasi-experimental design in which the performances of the learners in the three groups of I-DA, G-DA, and C-DA were compared. A quasi-experimental design is a design in which nearly all the elements of a true experimental design are present, and just one or two (usually through randomization) are missing. Due to the fact that choosing a random sample is practically impossible in the context where the study was conducted, the design of the study was quasi-experimental. Within this design, one control group and three experimental groups represented the independent variables of interactionist dynamic assessment, group dynamic assessment, and computerized dynamic assessment, and listening comprehension represented the dependent variable. The control variables of the study were gender, age, and proficiency level of the participants.

### *Participants*

To carry out the study, the Preliminary English Test (PET) was given to approximately 140 Iranian EFL learners, and 80 of them were selected as the sample of the study. The students were selected out of four English Language Institutes of Ahvaz, Iran, and their level of general English proficiency was intermediate. Their age range was between 18 and 32 years old. All the participants were females, and they were native speakers of Persian. The non-random availability sampling method was used to choose the participants in the study. The selected students were assigned to four equal groups: three experimental groups and one control group. One of the experimental groups was taught based on interactionist dynamic assessment ( $n = 20$ ), another experimental group was taught through group dynamic assessment ( $n = 20$ ), and the last experimental group was taught based on computerized dynamic assessment ( $n = 20$ ). The control group was taught on the basis of traditional instruction ( $n = 20$ ).

### *Instrumentation*

Three instruments were used in this study. The first instrument was the Preliminary English Test (PET). It was employed to assist the teacher in selecting the homogenous participants. According to this test, the learners who score between one standard deviation (SD) above and one SD below the mean were regarded as intermediate and were recognized as the target sample of the study. This test includes 60 items and was developed by Oxford University Press. It has 60 multiple-choice items, and based on it, the learners whose scores are 0 to 10 are beginners; the learners whose scores are 11 to 17 are known as a breakthrough; the learners whose scores are 18 to 28 are elementary; those learners whose scores are 32 to 39 are pre-intermediate; the students whose scores are 42 to 48 are intermediate; the learners whose scores are 48 to 54 are known as the advanced learners, and those whose scores are 55 to 60 are very advanced learners.

The second instrument was the listening section of a sample TOEFL Junior Standard Test which primarily serves to determine the proficiency level of the test takers. TOEFL Junior Standard test is a paper-based test consisting of 120 multiple-choice question items. Each section contains 45 four-choice question items with a total testing time of 2 hours. The Listening Comprehension section assessed the ability to listen and comprehend English with specific purposes, material purposes, and academic purposes. Two experts of the field were asked to give their ideas about the items included in the tests. The reliability of the abridged test was measured by using Cronbach's alpha. This listening comprehension test was used as the pretest and an equivalent test taken from another sample TOEFL Junior Standard Test was used as the post-test.

The third instrument consisted of animation-related listening comprehension quizzes. In fact, the materials which were used in this study comprised 18 authentic short animation videos. These were

used because video materials can stimulate students to listen to them, while providing a sample of authentic language use; similarly, they can help learners comprehend the different animation videos in which the language is used, can introduce the features of animation videos to the learners (Flowerdew and Miller, 2005), and can be entertaining and amusing. Here, the teacher makes animation-related listening comprehension quizzes used both as classroom materials and as tests in the dynamic assessment phase of the study.

#### *Procedures*

##### *Data collection procedures*

To do this research, 80 homogenous students were first selected and assigned to four equal groups of 20: three experimental groups and one control group. Then, the listening pretest was administered in order to measure the listening ability of the students before performing the treatment. In the first experimental group, i.e., I-DA group, each session the participants were asked to watch and listen to a short animation video and discuss and share their understanding in their own subgroups. Next, they were asked to join separately in the related listening comprehension quiz. Then, the teacher (the researcher) participated in each subgroup for a short period of time. The interaction between the assessor (the teacher acting as a competent peer) and the learner continued until the learner could achieve the correct answer.

In the second experimental group, the teacher asked each participant of the G-DA group in each session to answer the first question and ensured that other participants were actively listening by keeping a close eye on what they were doing. If the learner's answer was correct, the teacher asked her to discuss the answer and delve into why it was correct, and if the answer was incorrect, the teacher provided her with a correct form of mediation and instruction. The mediation was discussed between the assessor and the learner, and it was flexible. The teacher actually offered hints, leading questions, explicit feedback and recommendations which were in harmony with the interactionist group to DA, and which the participants followed. The procedures in the second experimental group were almost similar to those of the first experimental group. The only difference was that by using Ableeva's (2010) regulatory scale, the teacher gave the learner appropriate prompts and hints ranging from the most implicit to most explicit feedback until the learner could achieve the correct answer. The scale consisted of 10 levels of intervention. If the implicit end of the scale was unsuccessful, the assessor used a more explicit form of mediation until the learner was able to make corrections. To the end, the assessor corrected the learners' mistakes. If they could not correct mistakes, the assessor provided students with careful and detailed definitions.

In the third experimental group, C-DA was implemented. C-DA can help the learners know their learning potential and the teachers will use the validated and reliable software developed by Pishghadam and Barabadi (2012); namely, Computerized Dynamic Assessment Test (CDAT). With regard to this software, it should be noted that it can easily be installed on any PC. The students have to enter some information such as their name, age, and major and after reading the software description, go directly into the passage and answer the questions / solve the task automatically. The learners can consult the assessor if a wrong answer appears. After completing the test, a scoring file is created on the desktop to know about the test taker's performance.

Moreover, the assessor used DIALANG assessment. It is an online assessment system which is dedicated to each learner who wants to achieve diagnostic information about their linguistic proficiency for only three of the main skills, such as Reading, Listening, and Writing. DIALANG's Assessment Framework and self-assessment statements are based on the Common European Framework of Reference (CEFR) for Languages; therefore, it also gives feedback on the diagnostic (the strength and the weakness) of the learner's proficiency and advices on how to develop language skills. To assess the learners' proficiency level in this research, they were asked to visit the DIALANG site on the internet. Because each test took about two hours for students to complete, the whole process lasted two days until the results of the all-learners' proficiency level was determined.

In the control group, during each session, the teacher asked the participants to listen to the short animation video and share their understanding in their own groups, and then separately answer the test items. In this group, the teacher did not join the groups to interact or mediate their performance in listening comprehension test items. This means that the teacher did not use GDA procedures. Instead, she gathered the learners' test papers and announced their scores in the following session. After the pretest and treatment sessions, the teacher gave the listening comprehension post-test to the learners of the experimental groups and the control group. This procedure followed until the last session. The whole treatment lasted 15 sessions of 45 minutes. In the first and the second sessions, the PET and pretest of listening was administered.

#### *Data Analysis procedures*

In order to analyze the collected data and to run parametric tests, IBM SPSS Statistics 23 was used. First, the Shapiro-Wilk test was employed to measure the exact normality of the data. Second, to detect the developments of each group from pretest to post-test, a paired-samples *t*-test was administered. In addition, to compare the listening comprehension development of the learners in different groups, one-way ANCOVA was run.

## Results

### Preliminary Analyses

Before running the parametric statistics required in this study (i.e., *t* test and ANCOVA), assumptions underlying them had to be checked. One of the most important assumptions was the assumption of normality, the results of which are displayed in Table 1 below:

**Table 1-** Results of the Kolmogorov-Smirnov's Test of Normality

Groups	Tests	Kolmogorov-Smirnov		
		Statistic	<i>Df</i>	Sig.
IDA	Proficiency Test	.186	20	.068
	Pretest	.150	20	.200
	Post-test	.192	20	.052
GDA	Proficiency Test	.176	20	.104
	Pretest	.150	20	.200
	Post-test	.163	20	.172
CDA	Proficiency Test	.162	20	.178
	Pretest	.163	20	.174
	Post-test	.160	20	.193
CG	Proficiency Test	.168	20	.140
	Pretest	.166	20	.150
	Post-test	.188	20	.062

In the above table, the *p* values under the Sig. column should be compared with the .05 level of significance, and a *p* value greater than .05 indicates no violation of the assumption of normality. As it could be noticed in Table 1, all the *p* values lined up under the Sig. column exceed .05, which means

that for all the groups of IDA, GDA, CDA, and CG, the OQPT scores as well as pretest and post-test scores formed normal distributions. In addition to the assumption of normality, other required assumptions such as homogeneity of variance, linearity, and homogeneity of regression slopes were checked and no violations of these assumptions were ensured.

### Comparison of the Proficiency Test Scores

As it was mentioned before, those 80 intermediate learners were chosen as the participants of the study, and formed the IDA, GDA, CDA, and CG groups. To further ensure the homogeneity of the learners after they were assigned to these four groups, their OQPT test scores were compared through an independent-samples *t* test. The obtained results are shown in Tables 2 and 3:

**Table 2 - Descriptive Statistics for the OQPT**

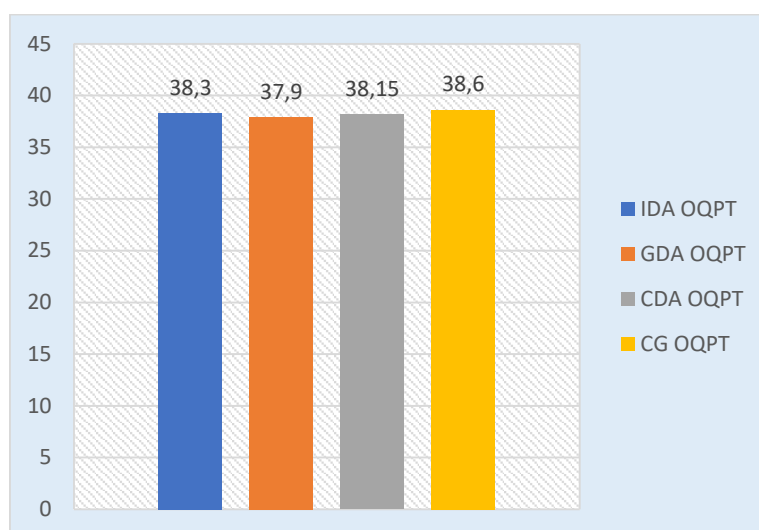
	<i>N</i>	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
					Lower Bound	Upper Bound
IDA	20	38.30	4.34	.97	36.26	40.33
GDA	20	37.90	4.93	1.10	35.59	40.20
CDA	20	38.15	4.96	1.11	35.82	40.47
CG	20	38.60	3.93	.88	36.75	40.44
Total	80	38.23	4.48	.50	37.23	39.23

The IDA, GDA, CDA, and CG learners' mean scores on the placement test were 38.30, 37.90, 38.15, and 38.60, respectively. In order to determine whether the differences among these mean scores (and thus these four groups) on the OQPT was statistically significant or not, the researcher had to examine the *p* value under the Sig. column in the one-way ANOVA table (Table 3). A *p* value less than .05 would suggest a statistically significant difference among these groups, and a *p* value larger than .05, on the other hand, would indicate a difference which failed to reach statistical significance.

**Table 3 - Results of One-way ANOVA Comparing the OQPT Scores of the Learners**

	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	Sig.
Between Groups	5.138	3	1.713	.082	.970
Within Groups	1583.350	76	20.834		
Total	1588.488	79			

Table 3 shows that there was not a statistically significant difference in the OQPT scores of IDA ( $M = 38.30$ ,  $SD = 4.34$ ), GDA ( $M = 37.90$ ,  $SD = 4.93$ ), CDA ( $M = 38.15$ ,  $SD = 4.96$ ), and CG ( $M = 38.60$ ,  $SD = 3.93$ ),  $F(3, 76) = .082$ ,  $p = .97$  (two-tailed). This was so because the *p* value was found to be larger than the significance level ( $p > .05$ ). Hence, it could be inferred that the learners in the four groups were at approximately the same level of proficiency at the outset of the study. This approximate equality of the four groups' OQPT scores is also graphically represented in the bar graph in Figure 1 below:



**Figure 1.** OQPT mean scores of the IDA, GDA, CDA, and CG learners

Figure 1 clearly illustrates that there were very small, negligible differences among the OQPT scores of the IDA, GDA, CDA, and CG learners. Consequently, the four groups of learners were at roughly the same level of proficiency before the experiment commenced.

#### ***Effects of IDA on Listening Comprehension***

The first research question of the study was: Does using interactionist dynamic assessment have any significant effects on Iranian EFL learners' listening comprehension? To answer this research question, the listening comprehension pretest scores of the learners in the IDA and CG groups had to be compared, together with the subsequent listening comprehension post-test scores. This could be achieved through two separate independent-samples *t* tests (one for the pretest analysis and one for the post-test analysis). However, to control for any possible differences between the IDA and CG prior to the treatment, one-way ANCOVA was conducted. This way the researcher could control for any possible differences between the two groups on the pretest and then compare their (adjusted) post-test scores. The results of the ANCOVA test are presented below:

**Table 4 - Descriptive Statistics for Listening Comprehension: Post-test Scores of the Learners in the IAD and CG**

Groups	Mean	Std. Deviation	<i>N</i>
IDA	27.95	2.48	20
CG	25.30	3.11	20
Total	26.62	3.08	40

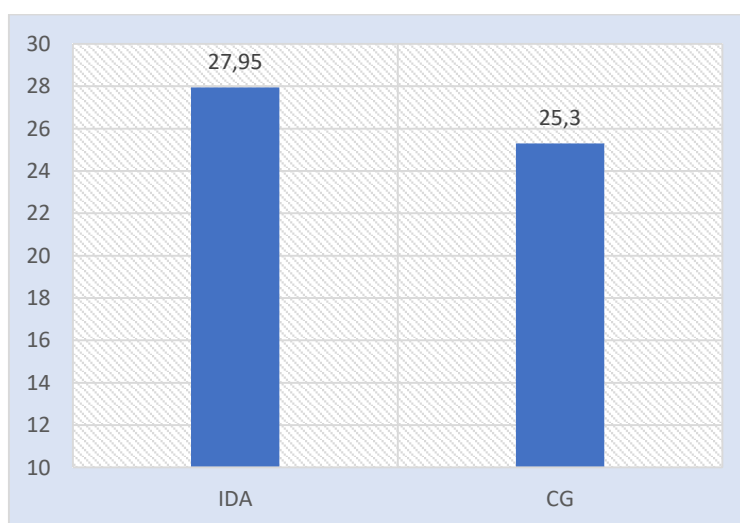
Such descriptive statistics as mean and standard deviation are shown for both IDA and CG learners in Table 4. The listening comprehension post-test mean score of the CG ( $M = 25.30$ ) was less than the listening comprehension post-test mean score of the IDA ( $M = 27.95$ ). To determine whether this

difference was a statistically significant one or not, the *Sig* (2-tailed) column in the ANCOVA table below should be observed:

**Table 5 - Results of One-Way ANCOVA for Post-test Scores of the Learners in the IAD and CG**

Source	Type III Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>Sig.</i>	Partial Eta Squared
Corrected Model	349.49	2	174.74	295.47	.00	.94
Intercept	12.49	1	12.49	21.11	.00	.36
Pretest	279.26	1	279.26	472.20	.00	.92
Groups	83.80	1	83.80	141.69	.00	.79
Error	21.88	37	.59			
Total	28727.00	40				
Corrected Total	371.37	39				

In Table 5, if you find Groups in the leftmost column and read across this row, under the *Sig.* column, you can find the *p* value, which should be compared against the significance level. The *p* value here was smaller than the specified level of significance ( $.000 < .05$ ), indicating that there was a statistically significant difference between the IDA and CG learners' listening comprehension post-test scores. This means that the treatment (i.e. interactionist dynamic assessment) significantly and positively affected the listening comprehension of the intermediate EFL learners in the IDA. Under Partial Eta Squared, the relevant value was .79, which shows that being in different groups (IDA vs. CG) accounted for 79% of the variance in the post-test scores of the learners. The results obtained in this part are shown in Figure 2 below:



**Figure 2.** Listening comprehension post-test mean scores of the IDA and CG learners



It is clear in Figure 2 that there was a considerable difference between the post-test scores of IDA and CG learners, with the former outperforming the latter. Consequently, it could be argued that interactionist dynamic assessment positively and significantly influenced intermediate Iranian EFL learners' listening comprehension.

#### ***Effects of GDA on Listening Comprehension***

The second research question of the study was: Does using group dynamic assessment have any significant effects on Iranian EFL learners' listening comprehension? As it was done with the preceding research question, one-way ANCOVA was conducted to capture the possible differences between the listening comprehension post-test scores of the learners in the GDA and CG. Tables 6 and 7 present the obtained results.

**Table 6 - Descriptive Statistics for Comparing the Post-test Scores of the Learners in the GDA and CG**

Groups	Mean	Std. Deviation	<i>N</i>
GDA	29.90	3.07	20
CG	25.30	3.11	20
Total	27.60	3.84	40

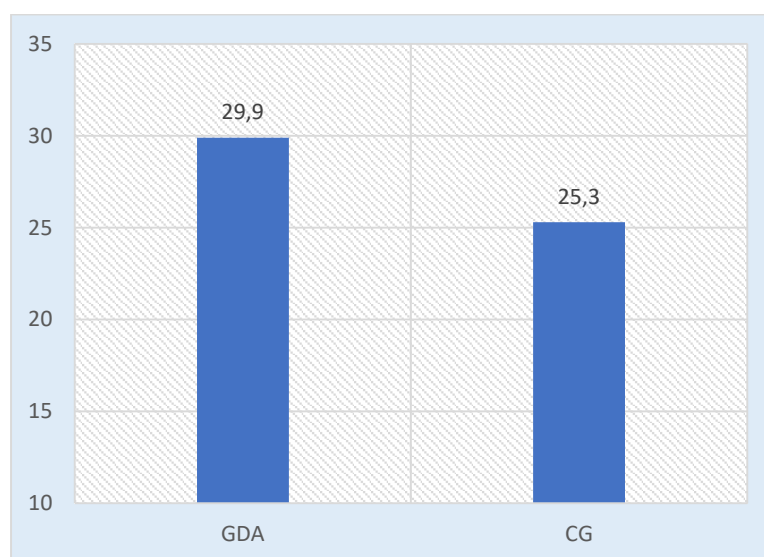
Table 6 shows that the post-test mean score of the CG ( $M = 25.30$ ) was less than that of the GDA ( $M = 29.90$ ). To find out whether this difference in the post-test scores of the GDA and CG learners was a significant one or not, one had to look down the *Sig.* (2-tailed) column in front of Groups in Table 7:

**Table 7 - Results of One-Way ANCOVA for Comparing the Post-test Scores of the Learners in the GDA and CG**

Source	Type III Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>Sig.</i>	Partial Eta Squared
Corrected Model	548.85	2	274.42	379.58	.00	.95
Intercept	20.45	1	20.45	28.29	.00	.43
Pretest	337.25	1	337.25	466.48	.00	.92
Groups	234.77	1	234.77	324.73	.00	.89
Error	26.75	37	.72			
Total	31046.00	40				
Corrected Total	575.60	39				

In Table 7, in front of Groups, under the *Sig.* column, the *p* value was smaller than the specified level of significance ( $.000 < .05$ ), indicating that the treatment (group dynamic assessment) was effective in improving the GDA learners' listening comprehension. Under Partial Eta Squared, the corresponding value was .89, which shows that the treatment accounted for 89% of the variance in the

listening comprehension post-test scores of the GDA and CG learners. The results obtained for this part are shown in Figure 3.



**Figure 3.** Post-test mean scores of the GDA and CG learners

It is clearly seen in Figure 3 that the difference between the post-test scores of the GDA and CG learners was substantial; in consequence, it could be inferred that group dynamic assessment positively (and significantly) influenced intermediate Iranian EFL learners' listening comprehension.

#### ***Effects of CDA on Listening Comprehension***

The third research question of the study was: Does using computerized dynamic assessment have any significant effects on Iranian EFL learners' listening comprehension? Again one-way ANCOVA was run to find out the possible differences between the listening comprehension post-test scores of the learners in the CDA and CG. The results are demonstrated in Tables 8 and 9.

**Table 8 - Descriptive Statistics for Comparing the Post-test Scores of the Learners in the CDA and CG**

Groups	Mean	Std. Deviation	<i>N</i>
CDA	28.50	3.31	20
CG	25.30	3.11	20
Total	26.90	3.56	40

It is evident in Table 8 that the post-test mean score of the CDA ( $M = 28.50$ ) was larger than the post-test mean score of the CG ( $M = 25.30$ ). To figure out whether this difference between the two mean scores could reach statistical significance or not, the researcher had to examine the  $p$  value under the Sig. (2-tailed) column in front of Groups in the following one-way ANCOVA table below:

**Table 9 - Results of One-Way ANCOVA for Comparing the Post-test Scores of the Learners in the CDA and CG**

Source	Type III Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>Sig.</i>	Partial Eta Squared
Corrected Model	478.25	2	239.12	510.13	.000	.96

Intercept	11.07	1	11.07	23.63	.000	.39
Pretest	375.85	1	375.85	801.82	.000	.95
Groups	112.49	1	112.49	239.98	.000	.86
Error	17.34	37	.46			
Total	29440.00	40				
Corrected Total	495.60	39				

As it could be seen in Table 9, there was a statistically significant difference between the post-test mean scores of the CDA and CG learners,  $F(1, 37) = 239.98, p = .000 < .05$ . This implies that the treatment (computerized dynamic assessment) was effective in improving the CDA learners' listening comprehension. The magnitude of this difference, shown under the Partial Eta Squared column, was found to be .86, indicating that the treatment accounted for 86% of the variance in the listening comprehension post-test scores of the CDA and CG learners. The results obtained above are depicted in the bar graph in Figure 4 below:

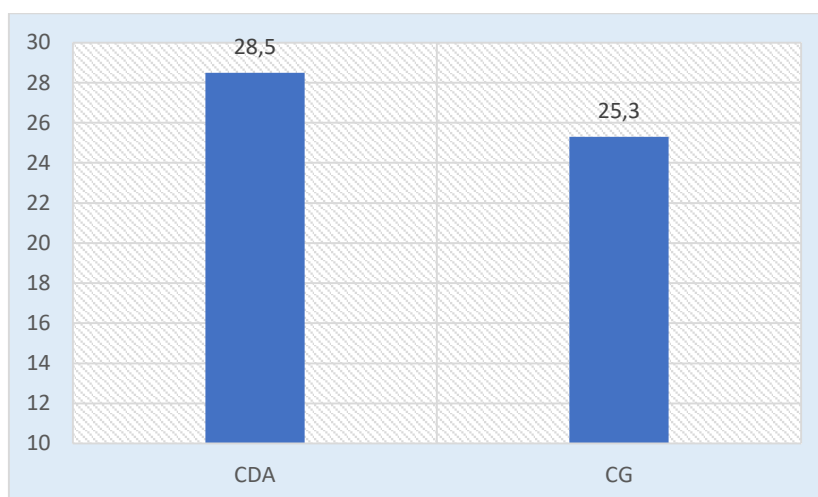


Figure 4. Post-test mean scores of the CDA and CG learners

The bar graph in Figure 4 demonstrates that there was a considerable difference between the post-test scores of the CDA and CG learners; accordingly, it could be concluded that computerized dynamic assessment positively (and considerably) affected intermediate Iranian EFL learners' listening comprehension.

**Comparing the Effects of IDA, GDA, and CDA**

The last research question of the study asked was: Are there any significant differences among the effects of interactionist dynamic assessment, group dynamic assessment, and computerized dynamic assessment on Iranian EFL learners' listening comprehension? To find an answer to this, another one-way ANCOVA was employed to find out the possible differences among the listening comprehension post-test scores of the learners in the IDA, GDA, and CDA. The descriptive statistics results are displayed in Tables 10.

Table 10 - Descriptive Statistics for Comparing the Post-test Scores of the Learners in the IDA, GDA, and CDA

Groups	Mean	Std. Deviation	N
IDA	27.95	2.48	20
GDA	29.90	3.07	20

CDA	28.50	3.31	20
Total	28.78	3.04	60

Table 10 reveals that there were differences among the post-test mean score of the learners in IDA ( $M = 27.95$ ), GDA ( $M = 29.90$ ), and CDA ( $M = 28.50$ ). In order to find out whether these differences among these three mean scores were of statistical significance or not, the relevant  $p$  value in the one-way ANCOVA table were checked (Table 11).

**Table 11 - Results of One-Way ANCOVA for Comparing the Post-test Scores of the Learners in the IDA, GDA, and CDA**

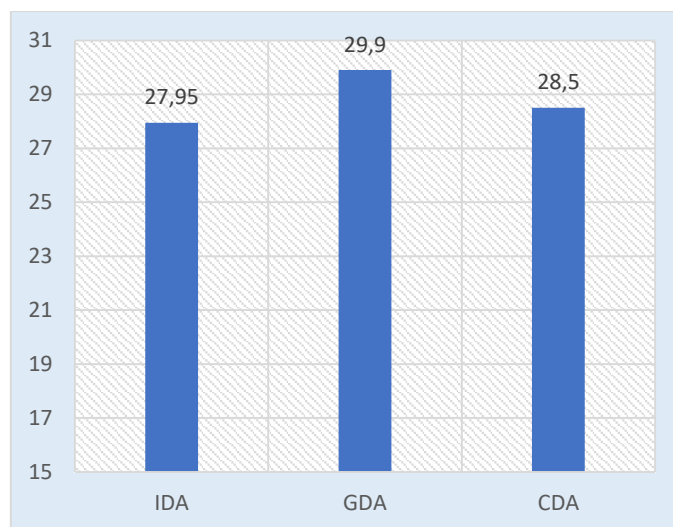
Source	Type III Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>Sig.</i>	Partial Eta Squared
Corrected Model	506.76	3	168.92	239.94	.000	.92
Intercept	49.30	1	49.30	70.04	.000	.55
Pretest	466.32	1	466.32	662.40	.000	.92
Groups	41.67	2	20.83	29.60	.000	.51
Error	39.42	56	.70			
Total	50255.00	60				
Corrected Total	546.18	59				

It is clearly seen in Table 11 that there was a statistically significant difference among the post-test mean scores of the IDA, GDA, and CDA learners,  $F(2, 56) = 29.60$ ,  $p = .000 < .05$ . This indicates that using different methods of dynamic assessment differently affected EFL learners' listening comprehension. The effect size for this difference, shown under the rightmost column of the table, was found to be .51, that is a very large effect size. The exact locations of the differences among these three groups are pinpointed in the post hoc test table below:

**Table 12 - Results of Bonferroni Post Hoc Test for Comparing the Post-test Scores of the Learners in the IDA, GDA, and CDA**

(I) Groups	(J) Groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
IDA	GDA	-1.95*	.26	.000	-2.60	-1.29
	CDA	-.45	.26	.283	-1.10	.20
GDA	IDA	1.95*	.26	.000	1.29	2.60
	CDA	1.49*	.26	.000	.84	2.15
CDA	IDA	.45	.26	.283	-.20	1.10
	GDA	-1.49*	.26	.000	-2.15	-.84

The  $p$  values presented in Table 12 show that GDA learners ( $M = 29.90$ ) performed significantly better than the CDA ( $M = 28.50$ ) and IDA ( $M = 27.95$ ) learners; however, the difference between the IDA and CDA learners failed to reach statistical significance ( $p = .283 > .05$ ). The results obtained above are demonstrated in the bar graph in Figure 5.



**Figure 5.** Post-test mean scores of the IDA, GDA, and CDA learners

Figure 5 demonstrates the fact that GDA learners significantly outperformed IDA and CDA learners, yet the difference between the post-test scores of the IDA and CDA learners was a slight one. The results presented above are discussed in the following section to come up with the findings of the study.

### Discussion

This section is allotted to the discussion of the main findings of the study with regard to the posed research questions.

#### ***The Effect of Interactionist Dynamic Assessment on Listening Comprehension***

Based on the results of the data analysis, interactionist dynamic assessment (IDA) positively and notably influenced Iranian intermediate EFL learners' listening comprehension. One plausible explanation for the remarkable effect of IDA on the learners' listening comprehension is, following Vygotsky's notion of cooperative interaction, the mediation within the course of exchanges between the examiner and the examinee. Through these, the examiner responds to the examinee's needs and continually reformulates his/her mediation. In the context of the present study, mediation, incorporating intentional and reciprocal interaction between the examiner and the learner in relation to the problems experienced by the learner and the developmental support given by the examiner, has allowed the examiner to collaborate on the listening task more closely with the learner, thereby enabling the assessor to move the learner to the next level of their ZPD.

Following the same lines of explanation and according to Vygotsky (1978), all our higher order thinking or mental activity is mediated by culturally-constructed artifacts, including language in social interactions. As a result of this mediation, higher mental functions such as memory, attention, rational thinking and development are promoted, and, consequently, perception of spoken language is fostered. In this process, the regulation of the shared activity is dynamic, i.e. shifting control gradually from the examiner to the learner and vice versa. Moreover, the participants, who have taken turns participating directly as primary interactants both among themselves and with the examiner, have benefited from each subsequent one-on-one exchange by building on earlier interactions that they have witnessed.

#### ***Effect of Group Dynamic Assessment on Listening Comprehension***

The results of the data analysis showed that group dynamic assessment (GDA) had a positive and meaningful impact on Iranian intermediate EFL learners' listening comprehension. A theoretical explanation for this observation could be that group DA, which is grounded in the sociocultural theory (SCT) by Vygotsky, may have enhanced learners' zone of proximal development (ZPD) in the context of the group. That is to say, exchanges initiated by the first interactant in the form of a question or comment can create an occasion for another's contribution. In the context of the present research, the participants' ZPD in the GDA group have been enhanced in the following way: If their answer was correct, they were asked to discuss the answer and delve into why it was correct, and if the answer was incorrect, the examiner provided them with a correct form of mediation and instruction. The overcoming of ZPD contributes to the transformation of their ability, listening comprehension in this case.

The findings of the study regarding the impact of GDA on listening comprehension could also be explained through the concept of development zone (DZ), proposed by Mercer (2000), who considered ZPD as being a dynamic rather than a static concept representing an individual's mental state at any given point. Such explanation takes account of the changing state of both the teacher's and learners' knowledge during the educational activity in the classroom. In other words, this zone is constantly resignified, as the teacher and their learners continue to engage in dialogue within a shared activity like discussion. This is in line with Holzman (2018), who interpreted ZPD as being actively and socially created rather than being an entity existing in psychological-cultural social space and time. Such view of ZPD as dynamic and socially constructed seems by nature more aligned with the purpose of DA.

#### ***Effect of Computerized Dynamic Assessment on Listening Comprehension***

The results of the data analysis proved that computerized dynamic assessment (CDA) positively and meaningfully influenced Iranian intermediate EFL learners' listening comprehension. One reasonable explanation for this observation could be envisaged from the perspective of cognitive psychology, according to which tailored mediation, which has occurred in the CDA group, could have contributed to the empowerment of the participants' listening comprehension through seriation thinking abilities, believed to be central to success in learning specific subjects including mathematics and logic. The decisive role of tailored mediation in CDA gains more importance when one considers the fact that test takers or examinees may be differentially affected by computer technology and, therefore, perform differentially on a computer-mediated task. This serves as the reason why the mediation provided could be optimally adapted to different examinees through computers, engaging them differentially in authentic contextualized language use tasks.

In line with Tzuril and Shamir (2002), the procedures in CDA provide more in-depth diagnoses of learner abilities and create more learning opportunities when teachers are present than when mediation is provided exclusively by the computer. Consequently, the findings of the study regarding the influence of CDA on listening comprehension could also be analyzed through the lens of regulation, i.e. the mutual control of the examiner and the examinee on the proceeding of the interaction, which includes the two modes of other-regulation and self-regulation. Other-regulation, involving explicit or implicit mediation (i.e., varying levels of assistance) by a capable peer or teacher, was realized through the examiner's providing hints, asking questions, etc. while working on the examinee's ZPD, thereby contributing to his improved attendance to the aural input. On the other hand, self-regulation, as the learner's ability to perform an activity without or with only minimal support from the examiner or a capable peer, has occurred in the context of dynamic assessment practice through computer software in the CDA group.

The findings of the study in this respect coincide with those of the Iranian researchers Mehri Kamrood, A., Davoudi, M., Ghaniabadi, S., & Amirian, S. M. R. (2019) who designed and implemented an online computerized dynamic test of pragmatic knowledge of Iranian EFL learners in an attempt to investigate how a DA procedure could shed more light on the unaccounted areas of EFL learners' PC abilities and found that due to test-takers' different Zone of Proximal Development (ZPD)

levels, their level of responsiveness to mediation was significantly different from one to another. In the context of the present study, this was also observed and named tailored mediation, which accounted for the development of listening comprehension in the participants. The results of the present study are also consistent with those of Anton (2009), Shabani (2012) and Mehri Kamrood et al., (2019) who confirm the positive effects of using computerized dynamic assessment in EFL contexts.

### ***The Differential Effect of Interactionist, Group, and Computerized Dynamic Assessment on Listening Comprehension***

According to the results, there was a significant difference between the IDA and CDA groups on the one hand and the GDA group on the other. That is to say, the participants in the former groups were outperformed by those in the latter one. However, no significant difference was observed between the IDA and CDA groups regarding their scores on the listening comprehension post-test. In fact, while there was no statistically meaningful difference between the IDA and CDA groups' listening comprehension scores, the GDA group differentially affected the participants' listening comprehension performance. Generally, the findings of the study accord with those of Alavi and Taghizadeh (2014), Antón (2009), Ebadi and Rahimi (2019), who found that the provision of mediation can boost participants' learning potential in the areas focused on and that DA has positive effects on the learners.

### **Conclusions**

The present study was designed to shed light on our understanding of the relative effect of three types of dynamic assessment; namely, interactionist, group, and computerized, on the listening comprehension of Iranian EFL learners. Based on the above-mentioned findings, it is worth mentioning that although caution should be exercised in generalizing this research outcome to the non-Iranian language learners, the fact that these studies strengthened the positive role of dynamic assessment in listening comprehension development would highlight the general effectiveness of dynamic assessment in language learning. Not only did the results of the present study highlight that dynamic assessment is effective in language learning, but they also made it evident that theories and principles behind dynamic assessment are valid in promoting learning. Accordingly, teachers and curriculum developers may creatively use the theories and principles behind dynamic assessment for the purpose of language instruction, including the receptive skill of listening comprehension.

The findings of the study suggest an alteration in the traditional models of listening comprehension assessment which emphasizes psychometric quantification of students' performances and offer no opportunities for learner-teacher interaction. Specifically, language teachers can make use of online CDA procedures both inside and outside the classroom contexts as a means for fulfilling a major goal, i.e. they can make use of the results of such procedures for diagnostic purposes. Actually, before the introduction of CDA into the field of diagnostic assessment, teachers could only make use of learners' level of independent performance or ZAD (i.e., actual scores), yet DA equips them with more diagnostic tools such as mediated and learning potential scores as well as the learners' scoring profiles generated right after they finished the test. Teachers could use these diagnostic tools for the whole class or for individual learners.

### **References**

- 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.
- Alavi, S. M., & Taghizadeh, M. (2014). Dynamic assessment of writing: The impact of implicit/explicit mediations on L2 learners' internalization of writing skills and strategies. *Educational assessment*, 19(1), 1-16.
- Alderson, J. C., Percsich, R., & Szabo, G. (2000). Sequencing as an item type. *Language Testing*, 17(4), 423-447.

- Alonazi, Z. (2017), Examining validity in computerized dynamic assessment. *Explorations in English Language and Linguistics*, 5 (1), 55-70.
- Anton, M. (2009). Dynamic assessment of advanced second language learners. *Foreign Language Annals*, 42(3), 576-598.
- Ebadi, S., & Rahimi, M. (2019). Mediating EFL learners' academic writing skills in online dynamic assessment using Google Docs. *Computer Assisted Language Learning*, 32(5-6), 527-555.
- Grigorenko, E. L. (2009). Dynamic assessment and response to intervention: Two sides of one coin. *Journal of Learning Disabilities*, 42, 111-132.
- Holzman, L. (2018). Zones of proximal development: Mundane and magical. In J. P. Lantolf, M. E. Poehner, & M. Swain (Eds.), *The Routledge handbook of sociocultural theory and second language development* (pp. 42-55). New York, NY, Routledge.
- Lantolf, J. P. & Poehner, M. E. (2007). Dynamic assessment of L2 development: Bringing the past into the future. *Journal of Applied Linguistics and Professional Practice*, 1(1), 49-72.
- 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.
- Mehri Kamrood, A., Davoudi, M., Ghaniabadi, S., & Amirian, S. M. R. (2019). Diagnosing L2 learners' development through online computerized dynamic assessment. *Computer Assisted Language Learning*, 6, 1-30.
- Mercer, N. (2000). *Words and Minds. How We Use Words to Think Together*. London, Routledge.
- Pishghadam, R., & Barabadi, E. (2012). Constructing and validating computerized dynamic assessment of L2 reading comprehension. *Iranian Journal of Applied Linguistics*, 15(1) 73-95.
- Poehner, M. E. (2009). Group dynamic assessment: Mediation for the L2 classroom. *TESOL Quarterly*, 43(3), 471-491.
- Shabani, K. (2012). Computerized dynamic assessment: an interventionist procedure to assess L2 reading. Paper presented at the *6th National and 3rd International Conference of E-Learning and E-Teaching*.
- Sternberg, Robert J., Elena L. Grigorenko (2002). *Dynamic Testing: The Nature and Measurement of Learning Potential*. Cambridge, UK, Cambridge University Press.
- Tzuriel, D., & Shamir, A. (2002). The effects of mediation in computer assisted dynamic assessment. *Journal of Computer Assisted Learning*, 18 (1), 21-32.
- Vygotsky, L. S. (1978). Socio-cultural theory. *Mind in society*, 6, 52-58.
- Vygotsky, L. ([1934] 1986). *Thought and language* (A. Kozulin, Trans.) Cambridge, MA., MIT.