Journal of Studies in Learning
and Teaching English
Volume. 5, Issue. 2, Ser. 10, (2016), 33-54

The Feasibility of Dynamic Assessment in EFL Classrooms: Evidence from an Iranian Context

Arshya Keyvanfar*
Islamic Azad University, Tehran North Branch

Mohammad Reza Darabi
Islamic Azad University, Tehran North Branch

Abstract. The present study aimed at investigating whether two types of assessment strategies, namely Dynamic Assessment and Non-Dynamic Assessment had any significant impact on elementary and pre-intermediate Iranian EFL learners' performance on their final achievement test including reading comprehension, vocabulary, grammar, and language functions. The participants of the study were 81 learners studying at the elementary and pre-intermediate levels in four classes in a language school in Tehran, Iran. Having tested them through the KET and PET respectively, one class at each level was randomly assigned as the experimental group. During 13 sessions, eight quizzes were administered in all the four classes. Following the principles of DA, participants of the two experimental groups received assistance when taking the quizzes while, in the two control groups, no assistance was provided. After the treatment, ANOVA was run on the results of the post-tests at the two levels, the results of which indicated that DA was only effective on the pre-intermediate learners and not on the elementary ones. The findings of this article will help English teachers with more effective teaching as they can set up an atmosphere where leraners can actively participate in the process of assessment while at the same time learners can experience a different and dynamic learning environment which is enjoyable, constructive, creative and free from test anxiety.

Keywords: ZPD, dynamic assessment, formative assessment, teacessor

Received: February 2016; Accepted: May 2016
*Corresponding author

1. Introduction

In today's world, assessment has become a fact of everyday life since people are assessed in order to obtain a driver's license, obtain a diploma, gain admittance to a university, earn a promotion at work, or receive credit for completing a course of study. Nowadays, even in the most modern systems of education, when people are assessed, great effort is usually made to assure that the assessment procedure is the same for everyone, and any interaction or assistance during the assessment itself is seen as inappropriate and unfair or even sometimes cheating.

EFL teachers have often argued and complained about the quality of various standardized tests and the effects of these tests on our teaching and learning outcomes. Indeed, we often feel helpless when confronted with large-scale testing practices. To us, assessment activities are easily distinguishable from other classroom activities because they are not integrated into our teaching process. In other words, assessment activities do stand out as different because they are usually formal and threatening. These activities are different from ordinary classroom activities in that they provide a way of assessing the students' performances and giving feedback in the form of a score or other information (e.g., notes in the margin, oral criticisms, or teacher explanations) that can enlighten the students and teachers about the effectiveness of the learning and teaching involved.

Another function of assessment is to offer information about how students are learning and teachers are teaching to improve the results. Traditional testing, as an instrument, has been quite product-based, and it has failed to provide process-based evidence of the given instruction. Based on the Common European Framework of Reference for Language (2011), the problems associated with traditional testing have also masked what the students really know or, in the case of English as a second/foreign language, what the students can do. As Aninao, Padella, and Sung (1996) found, traditional assessment has emphasized the measurement of a given body of defined and discrete knowledge as determined by a student's performance on a single test. This approach has often been limited to the assessment of students' achievement at a spe-

cific point in time and has provided little information about teaching and learning processes.

Grigorenko and Sternberg (1998) and Sternberg and Grigorenko (2002) have advocated the use of dynamic testing procedures instead of non-dynamic testing to measure individuals' strengths and weaknesses in cognitive skills, as well as their learning potential. Tests are usually non-dynamic as they require an examiner to give the test to the examinee without providing any feedback about their performance during the test. Dynamic tests, however, can be used to promote learning during the assessment by giving meaningful feedback about test performance to the examinees.

At this point, dynamic assessment in language learning, which applies Vygotsky socio-cultural theory into assessment, might offer new insights to assessment in language classrooms. Before Vygotsky, the general view about assessment was that the independent problem-solving was the only valid indication of mental functioning, whileVygotsky (1978) argued against this view by suggesting that independent problem-solving reveals only a part of learners' mental ability, that is their actual developmental level. A learner's potential developmental level is as important as the actual developmental one. Vygotsky discussed that responsiveness to assistance is an indispensable feature for understanding cognitive ability because it provides an insight into the person's future (potential) development. Hence, as Day (1993) suggested, it can be said that looking at the immediate performance of a learner and deciding on the achievement of any particular learner not only gives us a partial picture of that learner's performance, it also misleads us about our future analysis and planning based on the results of the current test.

This research article took an in-depth look at the issue of dynamic assessment from the standpoint of Vygotsky's Theory on Zone of Proximal Development (ZPD) in ELT class. In so doing, it investigated whether two types of assessment strategies, namely Dynamic Assessment and Non-Dynamic Assessment had different impacts on elementary and pre-intermediate Iranian EFL learners' performance on their final achievement test including reading comprehension, vocabulary, grammar, and language functions.

2. Review of the Current Literature

Product-oriented testing is one of the most widely supported evaluation methods in today's educational settings, including second and foreign language learning contexts. Many language teachers around the world still use final tests in their curricula. As Yildirim (2008) stated the basic idea behind testing students after instructing them for a certain amount of time is to understand how much they have progressed in the subject being taught. On the other hand, as he continued, it is not a rare case to hear a teacher saying that s/he does not understand why some students do very well in the class but cannot gain high grades in the tests.

Vygotsky (1978) believed that the normal learning situation for a learner was a socially meaningful cooperative activity. He viewed non-dynamic or traditional tests at schools which only look at a learner's individual problem-solving skills as insufficient, and proposed that the progress in concept formation achieved by the learner in cooperation with a more mature and knowledgeable counterpart (an adult, a teacher, or a more competent peer) would be a much more practical and logical way to look at the capabilities and readiness of learners. Vygotsky (1978) emphasized the role of social interaction and mediation or feedback in a learner's internalization.

According to Elliott, Lidz, and Shaughnessy (2004), the concept of DA is grounded in the assertion that conventional assessment is neither accurate nor instrumental in evaluating abilities and skills or satisfying the educational profiles of learners with special needs. Thus, the very initiation of DA is rooted in the concern that the information provided by traditional (conventional or non-dynamic) assessment is inaccurate with regard to learners' intellectual potential and of no real value to teachers. This assertion, in turn, is linked to the hypothesis that (a) children must be provided with a special type of evaluation, like DA; (b) this evaluation should result in more accurate data on the current level of performance and a profile of cognitive, motivational, and academic strengths and weaknesses; and (c) these data are translatable into remedial instructional strategies.

According to Poehner and Lantolf (2005), although DA has its ori-

gin in Vygotsky's concept of the ZPD, Vygotsky himself did not use the term DA when formulating his proposals on the cultural development of the individual nor when discussing his views on the importance of distinguishing between diagnostic and prognostic testing in educational settings. In fact, Vygotsky only talked about the range of possible interactive interventions to be used during ZPD assessment, such as asking leading questions, modeling, intending to do the assigned tasks and asking students to continue, and did not go further to produce any standardized procedure for the ZPD assessment. Nor did he make any particular distinction between ZPD assessment of general cognitive functions and ZPD assessment in content-based learning areas.

A number of isolated researchers had, in the early decades of the twentieth century, already experimented the concept of a modifiable intelligence (Lidz, 1992) including, among others, Selz (1935, as cited in Klauer, 2002); Ortar (1959), Haeussermann (1958) who is considered the mother of dynamic assessment, (Lidz & Elliott, 2000) and Schucman (1960, as cited in Lidz, 1992a, 2001). Kozulin (2005) stated that the concept of DA is only slightly younger than the concept of intelligence assessment itself, yet the amount of literature on DA is nowhere near that on conventional assessment. According to Grigorenko and Sternberg (1998), the disparity leads to the perception of DA as a "novel" assessment paradigm and its perception as an "adolescent" rather than a "mature adult" in the field of testing and assessment.

Although much investigation has been made at the theoretical level of dynamic assessment in language education, the number of empirical studies that could guide methodological applications is very limited. In one of these studies, Kozulin and Garb (2002) focused on dynamic assessment in second language reading achievement. They developed an instructional curriculum that included a dynamic assessment component focusing on helping learners develop general reading comprehension strategies. In the pre-test session of the study, the students were asked to read a simple passage in English and to answer a set of comprehension questions. In the mediation session, classroom teachers, who had been trained on mediation, reviewed the test with the students, mediating for them the reading strategies required in each item. For the

post-test purposes, following the mediation, students completed a post-test that closely paralleled the pre-test. Based on the analysis of the data, researchers devised a formula to calculate what they call 'Learning Potential Score' which quantified the gain between the pre- and post-tests. Then they argued that this score provided a complete picture of the students' ability than merely focusing on the final achievement tests in reading.

Still in another study, Anton (2003) focused on dynamic assessment as a placement procedure. He studied how dynamic assessment can be used to place students in a Spanish undergraduate language program. During the student placement oral exam, the examiner prompted students who had made some mistakes to give them the opportunity to revise their performance in appropriate ways. Students who could revise under prompting were considered to be more initiative than the students who could not, and all the students were placed in the program according to the results related to their response to mediation, or according to their potential for learning.

In a similar study, Poehner (2008) focused on dynamic assessment in advanced second language learning classrooms. In the study, first, learners constructed an oral narrative in the target language after watching a short video clip; they received no mediation in the first task. Then they were shown a second clip from the same story but this time they received hints, leading questions, suggestions, and explicit feedback when constructing their oral narratives. The assessment which focused on the performance differences between the first and second tests were used as the basis for an individualized instructional program in which participants were tutored in areas that had been identified during dynamic assessment as needing special attention. In other words, students' level of performance with the help of a mediator determined whether or not they need special attention in the language learning process.

Several scholars (e.g. Lantolf & Thorne, 2006; Sternberg & Grigorenko, 2002) have discussed the concept of dynamic assessment by differentiating it from non-dynamic assessment. They believe the goal of DA is usually to evaluate the students' learning ability and to help gain information useful for more effective instruction. NDA, however, is often

limited to measuring the students' current performance level. According to Sternberg and Grigorenko (2002), in NDA the examiner presents items, and the examinee is expected to respond to these items successively, without receiving any feedback or intervention. On the other hand, according to Sternberg and Grigorenko (2002, p. vii), DA is a procedure that takes the results of an intervention into consideration. During the intervention, the Examiner teaches the examinees how to perform better on individual items or the whole test (Sternberg & Grigorenko, 2002, p. vii). The final score is either the learning score representing the difference between pre-test (before learning) and post-test (after learning) scores, or the score on the post-test alone (Sternberg & Grigorenko, 2002, p. vii).

Finally, the review of the current literature revealed that dynamic assessment can be a useful framework to be used in language classrooms as it focuses on potential rather than final and actual achievement.

3. Method

Participants

Participants of this study were 81 general English elementary and pre-intermediate students within the age range of 18 to 30 at Tehran Institute of Technology Language Department. They were members of four intact classes, that is two intact classes for each level. The number of male and female learners was not equal (35 males and 46 females). All the participants' native language was Persian.

The pretest-posttest quasi-experimental design was used as for the purpose of the study. This study involved the independent variable of formative assessment varying in the two forms of dynamic and non-dynamic assessment as well as the mediating variable of language proficiency which was allowed to vary only at two levels of pre-intermediate and beginner. The dependent variable was considered to be the achievement scores of the learners in the four groups in their final exam.

Instrumentation

The first two instruments used in this study were KET and PET, which were administered for elementary and pre-intermediate groups respec-

tively to ensure the homogeneity of the two groups at each level. The second instrument used in this study consisted of eight short quizzes designed by the researchers based on the content of the selected units of the book which were administered through the course.

The second instrument used in this study was eight short quizzes (appendix A) which were administered through the course. For the elementary level, units 5, 6, 7, & 8 of English Result Textbook elementary level were covered and based on the same content eight quizzes were designed. That is, for each unit two quizzes were developed. The points covered in these included: countable and uncountable nouns; some/any; there is/there are; much, many, a lot of; imperatives; present continuous and present simple; order of adjectives describing colour, shape, age; how to ask for things, how to talk about food, about people's interest; free time, likes and dislikes; and how to invite or reply. In the pre-intermediate level, the units 1, 2, & 3 of English Result Textbook pre-intermediate level were covered and based on the same content eight quizzes were developed, i.e. for unit 1 two quizzes and for units 2 &3 three quizzes for each. The quizzes evaluated the language functions and skills such as reading and writing and language components like vocabulary and grammar while they were mostly based on the major topics including: possessives, adjective orders, comparative and superlative adjectives, past simple, like and would like, could (possibility), going to (predictions), and the topics such as: how to give and understand personal details, how to ask questions about people, how to compare weather in different places, how to suggest what to do, how to talk about what's going to happen, likes and dislikes in the covered units. There were about four to five sections and more than 25 items in each quiz.

The last instrumentations used for this study were the final exams designed for each level (Elementary and Pre-intermediate levels) which all the students took as the post-test. The post-tests included reading comprehension, language functions, grammar and vocabulary.

Procedure

Prior to the treatment, the KET and PET were administered for elementary and pre-intermediate levels respectively to ensure the homogeneity

of the groups. After running the pre-tests, one class at each level (elementary and pre-intermediate) was randomly assigned to the experimental group where dynamic assessment strategy was applied. The experimental groups at the elementary level included 21 participants while the control group had 20 participants. Also, at the pre-intermediate level, both experimental and control groups had 20 participants. Two teacessors were selected from among the experienced and qualified teachers who had gone through the essential courses including TESOL and CELTA in order to learn how to deal with different types of learners with different learning and cognitive styles. Each class consisted of 14 three-hour sessions with the post-tests run on the last session. The eight short quizzes were administered in all the four classes through the thirteen sessions. In the experimental groups, necessary feedback and mediation were given whereas in control groups no mediation was allowed. The quizzes for the two levels were designed based on the content of the selected units of the book and administered in two parallel classrooms with the same teacher, one with the teacessors' feedback and the other one without any feedback.

During the class sessions in the experimental group, the teacessor met the instructional objectives based on the level of the learners (i.e. elementary and intermediate). The teacessor also asked the students to do pair work and have interaction with each other to solve the problem. During the class session, the teacessor also was to provide the required feedbacks for the students based on the context of use. The following episodes show two examples from the sessions.

Table 1: Some examples from the class sessions (elementary level)

Level	Instructional Objective	the amount of feedback least → most	Ultimate Required Response
ELEMENTARY	some & any	He does not want …….. sugar in his coffee. (negative) (uncountable and negative)	He does not want any sugar in his coffee.
	adjective or adverb	Can you speak slow/slowly, please? (*speak* is a verb) (Use an adverb.)(speak slowly)	Can you speak slowly, please?
	present progressive tense	Listen. The teacher talks/'s talking. (It is now) (going on now!) ('s talking)	Listen. The teacher's talking.
	simple present tense	On Saturdays, our English class…….(start) at ten o'clock in the morning. (3rd person singular) (every Saturday!) (starts)	On Saturdays, our English class starts at ten o'clock in the morning.

These examples were provided to the elementary students. In the first case, the instructional objective was to instruct the uses of 'some' and 'any' in English grammar. The teacessor provided the instruction through assessing an example. The teacher also used mediation during the instruction by providing different hints to the students (such as mentioning that this is a negative statement and includes an uncountable noun, so the students would find that they should use 'any'). Then following examples are also taken from the pre-intermediate group.

Table 2: Some examples from the class sessions (pre-intermediate level)

Level	Instructional Objective	the amount of feedback least → most	Ultimate Required Response
Pre-intermediate	vocabulary	Big is the …… of small. (not synonym) (big #small) (opposite)	Big is the opposite of small.
	like+ …ing	We like swim/swimming in the pool at midnight! (like sth) (after *like* you should use ing forms!)	We like swimming in the pool at midnight!

For instance, the instructional objective is 'the use of ing'. The teacessor provided a statement assessing students' knowledge about the use

of 'ing'. Then the teacessor had to provide the one(s) who didn't understand the first statement with two forms of feedback and at the end the teacessor explained the use of the '-ing form' with a declarative statement, which is a kind of consciousness-raising.

The teacessor planned for the experimental group to receive the necessary feedback if they asked for some (clarification request). The teacessor even had to present the exact response necessary to solve the problem or answer the questions themselves. In other cases, the teacessor helped learners to guess the right answer that was actually what the teacessors preferred most of the time. To compare the efficiency of the two strategies in each of the groups, both groups at each level (elementary and pre-intermediate) were assessed through the same final test, and the results were analyzed. The post-tests included reading comprehension, vocabulary, grammar, and functions sections as all these were covered in both DA and NDA groups.

4. Results

Proficiency test results

To make sure that the two groups belonged to the same population, the researcher ran two independent samples *t*-tests on the pretest of each level (elementary and pre-intermediate). Tables 4.1 and 4.2 below show the descriptive statistics of the proficiency test results (the KET for the elementary and the PET for the pre-intermediate levels).

As it is shown in Table 3, the results of the skewness analysis obtained by dividing the statistic of skewness by the standard error of each of the two groups indicated that the assumption of normality was observed in the distribution of the scores of the two groups in the KET (-1.618 falls within the range of ± 1.96).

Table 3: Descriptive statistics, E level (KET)

N	Mean	Median	Std. deviation	Variance	Skewness
70	84.6463	90.0000	13.46674	181.353	-1.618

Also, the results of the skewness analysis obtained by dividing the statis-

tic of skewness by the standard error of each of the two groups (in the PET) indicated that the assumption of normality was observed in the distribution of the scores of the two groups.

Table 4: Descriptive statistics, PI level (PET)

N	Mean	Median	Std. deviation	Variance	Skewness
48	81.0500	82.0000	9.75324	95.126	-.687

Independent samples *t*-test, KET (E level) and PET (PI level) An independent samples *t*-test was run between the groups' scores on the KET to examine whether there was a significant difference between the two groups at the beginning of the study. As shown in Table 5, the Levene's test of homogeneity ($p = 0.228 > 0.05$) shows that equal variances were assumed and the criterion for running a *t*-test was met. Based on the results, the *t*-observed at 0.05 level of significance ($t = -0.061$, $df = 39$, $p = 0.952 > 0.05$) showed that there was no significant difference between the groups on the KET.

Table 5: Independent samples *t*-test, pretest (KET)

	Leven`s Test		**t-test for Equality of Means**				
	F	**Sig.**	**T**	**Df**	**Sig. (2-tailed)**	**Mean Difference**	**Standard Error Difference**
Equal Variances Assumed	1.5	0.228	-0.061	39	0.952	-0.1381	2.25948
			-0.061	37.905	0.951	-0.1381	2.24714

Another independent samples *t*-test was run between the groups' scores on the PET to examine whether there was a significant difference between the two groups at the beginning of the study. As shown in Table 6, the Levene's test of homogeneity ($p = 0.01 < 0.05$) shows that the means had to be compared with unequal variances. Based on the results, the *t*-observed at 0.05 level of significance ($t = -0.641$, $df = 38$, $p = 0.526 > 0.05$) showed that there was no significant difference between the groups on the PET.

Table 6: Independent samples *t*-test, pretest (PET)

	Leven`s Test		t-test for Equality of Means				
	F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Equal Variances	7.46	0.01	-0.641	38	0.526	-0.95	1.48302
Variances Not Assumed			-0.641	32.999	0.526	-0.95	1.48302

Independent samples *t*-test, KET and PET (various components)

As it is shown in Table 7, the Levene's test of homogeneity for Function, Grammar, and Vocabulary ($p = 0.56 > 0.05, p = 0.27 > 0.05$ and $p = 0.54 > 0.05$ respectively) shows that equal variances were assumed and the criterion for running a *t*-test was met. Based on the results, the t-observed at 0.05 level of significance ($t = -1.07, df = 33, p = 0.29 > 0.05$ for Function, $t = -1.37, df = 33, p = 0.18 > 0.05$ for Grammar and $t = -0.78, df = 33, p = 0.19 > 0.05$ for Vocabulary) showed that there was no significant difference between the groups on KET. At the same time, based on the results, the *t*-observed at 0.05 level of significance ($t = -1.21, df = 33, p = 0.23 > 0.05$ for reading and $t = -1.34, df = 33, p = 0.19 > 0.05$ for final score) showed that there was a significant difference between the groups on the KET.

Table 7: Independent samples *t*-test, KET (Various components)

	Leven`s Test		t-test for Equality of Means				
	F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Standard Error Difference
Equal Variances Assumed READING	10.45	0.00	-1.21 -1.42	33 27.05	0.23 0.17	-1.24 -1.24	1.02 0.87
Equal Variances Assumed FUNCTION	0.35	0.56	-1.07 -1.06	33.00 26.63	0.29 0.30	-0.40 -0.40	0.38 0.38
Equal Variances Assumed GRAMMAR	1.28	0.27	-1.37 -1.42	33.00 31.29	0.18 0.17	-0.64 -0.64	0.47 0.45
Equal Variances Assumed VOCABULARY	0.39	0.54	-0.78 -0.84	33.00 32.96	0.44 0.40	-0.26 -0.26	0.33 0.31
Equal Variances Assumed FINAL	5.57	0.02	-1.34 -1.51	33.00 31.42	0.19 0.14	-6.51 -6.51	4.85 4.30

As shown in Table 8, the Levene's test of homogeneity for Vocabulary ($p = 0.22 > 0.05$) shows that equal variances were assumed, and the criterion for running a *t*-test was met. Based on the results, the *t*-observed at 0.05 level of significance ($t = 1.84,\ df = 40,\ p = 0.07 > 0.05$) for Vocabulary showed that there was no significant difference between the groups on the PET. Similarly, based on the results, the *t*-observed at 0.05 level of significance ($t = 2.03,\ df = 40,\ p = 0.05$ for reading,$t = 2.83,\ df = 40,\ p = 0.01 < 0.05$ for function, $t = 2.83,\ df = 40,\ p = 0.01 < 0.05$ for Grammar and $t = 4.01,\ df = 40,\ p = 0.00 < 0.05$ for the final score) showed that there was a significant difference between the groups on the PET.

Table 8: Independent samples *t*-test, PET (Various components) components)

	Leven`s Test		**t-test for Equality of Means**				
	F	**Sig.**	**t**	**Df**	**Sig. (2-tailed)**	**Mean Difference**	**Standard Error Difference**
Equal Variances Assumed	0.01	0.92	2.03	40.00	0.05	1.29	0.63
READING			2.03	39.36	0.05	1.29	0.63
Equal Variances Assumed	2.00	0.16	2.83	40.00	0.01	0.76	0.27
FUNCTION			2.83	35.54	0.01	0.76	0.27
Equal Variances Assumed	3.46	0.07	5.59	40.00	0.00	1.48	0.26
GRAMMAR			5.59	38.76	0.00	1.48	0.26
Equal Variances Assumed	1.55	0.22	1.84	40.00	0.07	0.43	0.23
VOCABULARY			1.84	37.65	0.07	0.43	0.23
Equal Variances Assumed	2.93	0.09	4.01	40.00	0.00	10.12	2.52
FINAL			4.01	34.35	0.00	10.12	2.52

Estimating the difference between the means of components, posttest

The results of the test (Table 4.7) show that the F(1.465, 1.147, 1.871, 0.613 and 1.804 for reading, function, Grammar, Vocabulary and final scores of the E level students, respectively) at the alpha level of 0.05 is not significant and thus the null hypothesis is not rejected. In fact, there

is no significant difference between the pretest and posttest scores of E level students in the NDA and DA groups.

Table 9: ANOVA for D and ND groups (E-level)

ANOVA						
		Sum of Squares	Df	Mean Square	F	Sig.
Reading	Between Groups	12.876	1	12.876	1.465	.235
	Within Groups	290.095	33	8.791		
	Total	302.971	34			
Function	Between Groups	1.376	1	1.376	1.147	.292
	Within Groups	39.595	33	1.200		
	Total	40.971	34			
Grammar	Between Groups	3.471	1	3.471	1.871	.181
	Within Groups	61.214	33	1.855		
	Total	64.686	34			
Vocabulary	Between Groups	.576	1	.576	.613	.439
	Within Groups	31.024	33	.940		
	Total	31.600	34			
Final	Between Groups	356.201	1	356.201	1.804	.188
	Within Groups	6514.542	33	197.410		
	Total	6870.743	34			

The results of the test (Table 10) show that the F(4.140, 8.025, 31.303, 3.375 and 16.083 for reading, function, Grammar, Vocabulary and final scores of the PI level students, respectively) at the alpha level of 0.05 is significant and the null hypothesis is rejected. In fact, there is a significant difference between the pretest and posttest scores of PI level students in the NDA and DA groups.

Table 10: ANOVA for DA and NDA groups (PI-level)

ANOVA						
		Sum of Squares	Df	Mean Square	F	Sig.
Reading	Between Groups	17.357	1	17.357	4.140	.049
	Within Groups	167.714	40	4.193		
	Total	185.071	41			
Function	Between Groups	6.095	1	6.095	8.025	.007
	Within Groups	30.381	40	.760		
	Total	36.476	41			
Grammar	Between Groups	22.881	1	22.881	31.303	.000
	Within Groups	29.238	40	.731		
	Total	52.119	41			
Vocabulary	Between Groups	1.929	1	1.929	3.375	.074
	Within Groups	22.857	40	.571		
	Total	24.786	41			
Final	Between Groups	1074.643	1	1074.643	16.083	.000
	Within Groups	2672.769	40	66.819		
	Total	3747.412	41			

5. Discussion

The main objective of the present study was to examine whether DA could enhance Iranian EFL learners achievement scores. The answer signified that using DA could improve the learners' achievement in English. The study suggests that utilizing DA enhances students' motivation and motivational behavior to a great extent. The instructional value of this experimentation on EFL assessment lies in the fact that although DA is given as a group assessment, its results can be used for the development of individual learning plans for students with different learning needs. For example, students who demonstrated an average pretest performance, but insufficient learning potential should be taught learning and information-processing strategies to "learn how to learn". Students with an average pretest performance and high learning

potential should be given more challenging materials and more opportunity for independent studies. Students with low pretest performance and low learning potential need an intensive investment on their general learning and problem-solving skills with the help of very simple EFL materials. In other words, through DA, teachers can give learners enough self-confidence and motivation to learn more about their weak points while being tested.

There are some good reasons to use dynamic assessment in EFL classrooms. Dynamic assessment is not a viable alternative to traditional assessment. Some believe DA should not replace traditional assessment, but rather be used in conjunction with it (e.g., Lidz, 1987). DA may offer a less-biased measure of achievement for certain populations because it is less dependent on mainstream language skills and background experience (Sewell & Severson, 1974). It may especially be useful to differentiate various low-achieving students. Traditional tests are often subject to stress for low-achieving students. Items are scored "right" or "wrong" using an all-or-nothing mentality. DA, by contrast, gives multiple opportunities for success. Low-achieving students, therefore, can be differentiated along the continuum of how easily they learn. DA may inform instruction so that educational interventions can be more readily designed (Feuerstein, 1979; Haywood, 1992). If a test is susceptible to stress and students fail all items, we do not have useful data to gauge their academic functioning and plan appropriate interventions. DA has the potential to predict future student achievement because it attempts to measure the process and ease of learning. Presumably, those who learn with more ease will benefit more from classroom intervention and achieve at a higher level. As Embretson (2000) mentioned, the idea behind DA is to identify and assess changeable or remediable skills, describe and classify them, and select or devise instructional methods based on this classification. Ryba (1998) stated that DA was conceived as a package including both assessment and intervention: thus overcoming the gap between assessment and teaching. DA is neither an assessment instrument nor a method of assessing but a framework for conceptualizing teaching and assessment as an integrated activity of understanding learner abilities by actively supporting their development (Poehner, 2008). It is

also worth mentioning that the ideas of DA were often developed due to practical issues regarding assessing the potential of and developing pedagogical strategies for learners with special educational needs or inadequate educational experiences (e.g., new immigrants or the socially disadvantaged). Sternberg and Grigorenko (2001) argued that it was assumed that tests of abilities, when administered in a conventional, unassisted manner, assessed what the learner had already mastered.

They further argued that "true" abilities of such special learners, conceptualized as learning potential, are better captured through tests administered in an assisted manner. Thus, Grigorenko (2009) concluded that the driving motivation behind a variety of approaches to DA was to develop instruments capable of generating data indicative not of what a learner can do now, but rather what the learner can do in the future, assuming that pedagogical instructions are tailored to his or her needs.

Where as some DA supporters still advocate domain-general approaches (Feuerstein, et al., 2002), others argue for domain specificity of DA, stressing its importance for specific achievement areas including reading, math, and science (Camilleri, 2005; Hamers, Ruijssenaars, Sijtsma,1992). In spite of numerous investigations at the theoretical level of DA in language education, the number of empirical studies which could provide guidance for methodological applications is very limited. In one of the studies, Kozulin and Garb (2002) focused on dynamic assessment and second language reading achievement. DA has been pursued by the school and clinical psychologists as a method of more accurately assessing an individual's potential for future development by embedding instruction in the assessment process itself (Sternberg & Grigorenko, 2002).

The findings confirmed that DA had a positive impact on the achievement scores of the participants in the groups. To restate, the one-way ANOVA revealed that the learners had a gradual development in their ability as they received the treatment. However, as the statistical analyses showed, the experimental group gained a higher mean at each of the phases of measurement as compared to the control group.

In conclusion, the results indicated that DA was both feasible and effective in obtaining information on learners' learning potential. All

learners benefited, in varying degrees, from the mediation and were able to apply their acquired strategies to the new situations. Hence DA procedure also pin-pointed those learners who needed extra help with specific strategies. It was confirmed that learners with similar initial performance levels demonstrated different and in some cases dramatically different levels of ability to learn and use new learning strategies. The findings of this research confirmed the practical value of the EFL dynamic assessment procedure, showing that it provided in-depth information about different learning needs of the learners.

6. Conclusion

Overall, the results of the study suggested that at the pre-intermediate level, the group being instructed through DA strategy outperformed NDA group with a significant difference and thus the researchers could safely reject the null hypothesis of the study. However, at the Elementary level for the group being instructed through DA strategy the null hypothesis was not rejected. In fact, there was no significant difference between the pretest and post test scores of E level students in the NDA and DA groups and it might be due to the fact that Elementary learners were not still familiar with different forms of feedback given from the teacessor, another reason could be due to their lack of confidence and not being so autonomous learners.

DA is encouraging since the learners, and the teachers both enjoy the effect of education. DA enables students to take control of, and responsibility for their improvements, since they know exactly which strategies and information they should apply. It also enables the teacessor to supply, if necessary, additional material to those students whose improvement was limited. In addition, since the quality of mediation is the primary component in DA, it is recommended that we maintain a consistent standard of the mediation for the mediator to be used as guidelines while providing mediation for the learners. Thus, further research will be ideal if it can focus on designing consistent and high-quality mediation to be implemented in the dynamic assessment procedure.

English teachers can benefit from using DA strategy to teach more effectively. As most of the teachers face several difficulties in teaching

EFL to their students, using this strategy, they can provide a learning atmosphere in which students can actively participate in the process of instruction; moreover, learning will take place in a joyful and cooperative setting for both teachers and students. Similarly, learners can experience a different learning and assessment environment that creates a more constructive, creative, and autonomous testing/assessment environment and is free from test anxieties, bias and other sources of error. Also, test developers may include dynamic assessment principles in test construction and development to ease the implementation of real assessments. Materials developers may introduce a variety of practices and exercises in the lessons to establish more humanistic relationships and to use formative assessment results in their lessons. Last but not least, teacher educators can include this practice as a useful model in their teacher education programs.

References

Aninao, J., Padella, A., & Sung, H. (1996). Development and implementation of student portfolios in foreign language programs. *Foreign Language Annals, 29*(3), 429-438.

Antón, M. (2009). Dynamic assessment of advanced second language learners. *Foreign Language Annals, 42*(3), 576-598.

Camilleri, B. & Botting, N. (2013). Beyond static assessment of children's receptive vocabulary: the dynamic assessment of word learning (DAWL). *International Journal of Language & Communication Disorders, 48*(5), 565-581.

Day, J. D. & Cordón, L. A. (1993). Static and dynamic measures of ability: An experimental comparison. *Journal of Educational Psychology, 85*(1), 75.

Embretson, S. E. & McCollam, K. M. S. (2000). Psychometric approaches to understanding and measuring intelligence. In R.J. Sternberg (Ed.), *Handbook of intelligence* (pp. 423-444), NY: Cambridge University Press.

Feuerstein, R., Rand, Y. A., & Hoffman, M. B. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques.* Baltimore, MD: University Park Press.

Grigorenko, E. L. (2009). Dynamic Assessment and Response to Intervention Two Sides of One Coin. *Journal of Learning Disabilities, 42*(2), 111-132.

Grigorenko, E. L. & Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin, 124*(1), 75.

Hamers, J. H. M., Ruijssenaars, A., & Sijtsma, K. (1992). *Learning potential assessment.* Place: CRC Press.

Haywood, H. C. (1997). *Interactive assessment. Assessment of individuals with mental retardation.* San Diego: Singular Publishing Group.

Kozulin, A. & Garb, E. (2002). Dynamic assessment of EFL text comprehension. *School Psychology International, 23*(1), 112-127.

Lantolf, J. P., Thorne, S. L., & Poehner, M. E. (2000). *Sociocultural theory and second language development.* NY: Cambridge University Press.

Lidz, C. S. (1991). *Practitioner's guide to dynamic assessment.* NY: Guilford Press.

Lidz, C. S. E. (1987). *Dynamic assessment: An interactional approach to evaluating learning potential.* NY: Guilford Press.

Poehner, M. E. (2008). Dynamic assessment: *A Vygotskian approach to understanding and promoting L2 development* (Vol. 9), place: Springer Science & Business Media.

Ryba, K. (1998). Dynamic assessment and program planning for students with intellectual disabilities. *New Zealand Journal of Psychology, 27,* 3-10.

Sewell, T. E. & Severson, R. A. (1974). Learning ability and intelligence as cognitive predictors of achievement in first-grade black children. *Journal of Educational Psychology, 66*(6), 948.

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

Vygotsky, L. (1978). Interaction between learning and development. *Readings on the Development of Children, 23*(3), 34-41.

Yildirim, O. (2008). Vygotsky's sociocultural theory and dynamic assessment in language learning. *Anadolu University Journal of Social Sciences, 8*(1), 301-308.