Beyond Reading Comprehension: The Effect of Adding a Dynamic Assessment Component on EFL Reading Comprehension

Mehdi Mardani
University of Isfahn, Iran
Email: Mehdimardanii@yahoo.com

Manssour Tavakoli
University of Isfahan, Iran
Email: mr.tavakoli14@gmail.com

Abstract—Dynamic assessment (DA) stresses the need for unifying assessment and instruction. This paper presents an interactionist model of DA to assessment in reading comprehension of 30 Iranian male students who were selected based on available sampling procedure. Data collection procedures before and after implementation of DA were done through administration of multiple-choice reading comprehension test. The results of students' performance before and after implementation of DA were calculated through t-test. The results indicate significant improvement in student performance after implementation. Finally, the null hypothesis is rejected and it is concluded that incorporation of DA as a supplement procedure to standard testing has positive effective on both test performance and learning of students.

Index Terms—dynamic assessment, static assessment, reading comprehension

I. INTRODUCTION

Undoubtedly one of the most grueling and frustrating part of any educational course is the assessment part. Poehner (2008) asserts that "Students frequently echo this frustration when they are required to undergo regular assessment in order to demonstrate mastery of content or competency to pass to the next level of instruction" (p. 3). To make the situation even worse we raise the repetitious question of "why do we assess our students?" The fact is that assessment is usually looked at as an information-gathering tool (Bailey, 1996, cited in Poehner). Narrowing the function of assessment as only an information-gathering tool not only leads to bifurcation between teaching and assessment, but puts assessment in direct opposition to instruction. One possible way to combine the two distinct but related fields is the development of Dynamic Assessment (henceforth, DA). This reunification happens only when we integrate a mediation phase into our assessment (Lidz and Gindis, 2003). This view is in line with the approach to assessment and instruction described in the Sociocultural Theory of Mind (SCT), as developed by the Russian psychologist L.S. Vygotsky and his colleagues more than 80 years ago (Poehner, 2008).

According to Williams & Burden (1997) sociocultural theory of mind is part of a bigger paradigm called constructivism. Two names which are associated with constructivism are Piaget and Vygotsky. They differ mostly in the degree to which they value the role of social context in the development of language. For Piaget language develops as a result of gradual growth of general intellectual skills (Woolfolk, 2004). So, it can be said that Piaget theory is a developmental one. But, Lev Vygotsky offered an alternative to Piaget's stages of cognitive development. Vygotsky's Sociocultural Theory of Development is now a major influence in education (Woolfolk, A., 2004). Vygotsky advocates the primacy of social constructivist theory in which social interaction is the driving force in language development. Social constructivist theory, according to Nyikos and Hashimoto (1997), is mostly applied to address the learning through social interaction as delineated by the *zone of proximal development* which is the distance between a child's actual cognitive capacity and the level of potential development through mediation or scaffolding. So, under collaborative condition students reveal certain emergent functions which have not been yet fully internalized, or which have not been part of Zone of Actual Development (Kuzlin and Grab, 2002). These functions belong to ZPD. While static tests reveal information about the already existent abilities, we need a testing system which reveals information about the emergent abilities of learners. Hence, DA is proposed as a way of measuring the emergent abilities of students in the realm of ZPD.

Central to Vygotsky's theory was also the relationship between the development of thought and language. Vygotsky's theory views language first as social communication, gradually promoting both language itself and cognition. This assumption constitutes the base of social constructivist theory which emphasizes the importance of Socio-cultural factors in understanding what occurs in society and constructing knowledge based on interaction (Derry, 1999). It is

assumed that if we are to apply social constructivist theory to our EFL instruction setting, the evaluation of the students in these settings also needs a change and should be appropriately redesigned. According to Dixon-Krauss (1996), ZPD is a dynamic working model because it both guides and evolves through the social interaction that occurs during the learning activity. If this is the case, then we should think of a new model in order to measure the level of development of students in such dynamic context. The proposal for measurement in such contexts has been *dynamic assessment* (DA, hereafter). According to Heywood and Lidz (2007) DA is a recently developed, interactive approach to psycho educational assessment that follows a test-intervene-retest format, focuses on learning processes and modifiability, and provides the possibility of direct linkage between assessment and intervention.

The focus of this paper is on the implementation of DA in the assessment of reading comprehension in EFL classroom setting. In what follows, we will first discuss briefly the related review of literature, the concept of the ZPD and its realization in DA procedures; next, we will compare DA to Formative Assessment (henceforth, FA). Then, adopting an interactionist model, treatment is given and finally results are analyzed. The general goal of this study is to clearly show that DA, not as an alternative but as a supplement to standard test, can be administered in EFL classroom.

The present study

Based on what has been said so far it seems that doing further research on DA is a necessity. It is a valuable tool for those practitioners who really wish to fill the gap between teaching and testing, and link them together. By adopting a DA procedure students will no longer look at testing as something disgusting or frightening, rather they see it as another learning opportunity. This issue becomes considerably important when it comes to the context of Iran, where students are still tested quite improperly. That is, current measurement trend in Iran is the defective but prominent view of "teach to test". In a context like Iran, where no documented research on DA has been reported, conducting a research with the aim of paving the ground for further research seems justifiable. In this study the researcher aspired to bridge the gap between teaching and testing by adding a DA during the assessment of reading comprehension, hoping that it would have a significant effect on the performance of the group who receive support and help during mediation. For this purpose the following research question was proposed:

Is there a significant difference between the reading comprehension performance of dynamically-assessed students and non-dynamically-assessed ones?

Based on the above research question the following null hypothesis was proposed:

H. There is no significant difference between the reading comprehension performance of the two groups.

II. REVIEW OF LITERATURE

A. Theoretical Background

Although DA was the brainchild of Vygotsky's concept of ZPD, he himself did not use the term DA, (Poehner, & Lantolf, 2005). They believe that for first time it was A.R. Luria (1961), one of Vygotsky's most influential colleagues, who contrasts 'statistical' with 'dynamic' approaches to assessment. Unlike Standard tests which are psychometrically oriented, DA is rooted in the ZPD, so, its aim should be determining the extent to which the person's performance is modifiable. Because of the importance of ZPD, the author finds it essential to define ZPD in greater details.

1. Zone of Proximal Development

Vygotsky's sociocultural theory has two major features. The first, as it was mentioned is about the role of interaction in development. A second aspect of Vygotsky's theory is the idea that the capacity for cognitive development depends upon the "zone of proximal development" (ZPD). Vygotsky (1978) defines ZPD as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (p. 86). It, then, predicts the future development not *a priori* level. Poehner and Lantolf (2005) believe that "DA is very much in line with future-in-the-making models of development" (p. 237). As called for in Vygotsky's ZPD, assessment and instruction should be integrated as the means to move towards an always emergent (i.e., dynamic) future, rather than a fixed and stable steady state (ibid). Thus, it can be proposed that while static tests focus on the performance belonging to ZAD, DA tries to measure the emerging abilities which are not still fully developed and do not belong to the existing knowledge repertoire.

2. Dynamic Assessment versus Static Assessment

Emergence of new trends such as ZPD to language teaching calls not only for new approaches in language testing, but also for closer connection between these two fields. Static models of assessment which focus on student's existing knowledge and skills are no longer satisfactory. As Cioffi and Carney (1983) argue static measurement or non dynamic (NDA henceforth) procedures are the best choice when we want to evaluate the students' skill knowledge, but these procedures are insufficient for estimating the learning potential. Standardized testing can only inform much classroom-based assessment, even when the goal is to support learning (e.g., formative assessment, assessment for learning). Having said that, it should also be mentioned that whatever the shortcoming of static test were, they paved the ground for a more flexible approach to testing known as DA. DA development has been motivated by the inadequacy of standardized tests. The inadequacy can be summarized in the following points:

1). Static tests do not provide crucial information about learning processes

2). The manifested low performance level of many children, as revealed in standard test very frequently falls short of revealing their learning potential

3). In many static tests students are described mostly in relation to their relative position of their peer group, but they do not provide clear descriptions of the processes involved in learning, and

4). Static tests do not relate to non-intellective factors that can influence individuals' cognitive performance.

Dynamic assessment appeared not as a substitution for standardized test but as a complementary invention. It focuses on student's learning potential for learning. In other words, in the same way that ZPD focuses on the capacity and the level of potential development, DA focuses on learning potential and their ability to realize this potential during assessment process. It involves planned mediation of teaching and the assessment of effects of that teaching on subsequent performance (Campione & Brown, 1990). To summarize the differences so far, let us refer Sternberg and Grigorenko (2002), who make distinction between DA and NDA in three ways:

First, in terms of assessment goals, NDA focuses on 'products formed as a result of preexisting skills.' At the level of assessment administration, the non-dynamic paradigm does not permit 'feedback from examiner to test-taker regarding quality of performance' during the test procedure. Finally, with regard to the examiner's orientation in NDA, it is important 'to be as neutral and as uninvolved as possible toward the examinee (p. 28-29).

The difference between these two types of testing trends, or procedures does not stop here. Sternberg and Grigonernko (as cited in Lantolf and Thorne, 2006) identify three methodological and epistemological differences between DA and NDA. First, NDA focuses on past development while DA looks at the future development. The second issue is the relationship between examinee/examiner. In NDA examiner adopt a neutral role, but in DA examiner intervene in the assessment. Finally, in NDA there is no immediate feedback, but in DA feedback is given to students during a mediated procedure- and this is the crux of DA. According to Lantolf and Thorne (2006) "what makes a procedure dynamic or not is whether or not mediation is incorporated into the assessment process. In other words, fill-in-the-blank, multiple-choice, open-ended essay, or even oral proficiency test in themselves may or may not be dynamic" (p.331). Thus, there are no dynamic tests or instruments per se.

3. Approaches to Dynamic Assessment

DA is based on the Vygotskian notion of ZPD which captures the uniquely human potential to exceed beyond present capabilities in cooperation with others whose dialogic interaction mediates us to higher levels of functioning (Poehner, 2008). He adds that currently there a number of approaches and methods that fall under the umbrella term of term of DA. This is due to the fact that mediation can be implemented in a number of ways. But, According to (Lantolf and Poehner, 2004) there are two general approaches to DA, *interactionist* and *interventionist*.

Interactionist DA follows Vygotsky's tendency for dialogic interaction. In this approach, assistance emerges from the interaction between the mediator and the learner, and is therefore highly sensitive to the learner's ZPD. Imagine this scenario: a student is working on a reading comprehension multiple-choice item. He wisely identifies two incorrect options but when it comes to identifying the right answer between the other two option he gets stuck he cannot do that. Here you, as the mediator, steps in and tries to give hint. If the student cannot figure out the right answer more direct hints or support are provided until he gets to the right answer. Interventionist DA, on the other hand, remains closer to certain forms of static assessment and their concerns over the psychometric properties. Poehner (2008) states that the defining characteristic of interventionist DA is the use of standardized administration procedures and forms of assistance in order to produce easily quantifiable results that can be used to make comparisons between and within groups, and can be contrasted with other measures and used to make predictions about performance on future tests.

In addition to this dichotomous classification Sternberg and Grigorenko (2002) have described as *sandwich* and *cake* taxonomy of mediation. They believe that the sandwich format is much more in line with traditional experimental research designs in which treatment is administered following a pretest and a posttest. In this approach to DA, a mediation phase is similarly "sandwiched" between pretest and posttest that are administered in a non-dynamic manner. The *cake* format, however, refers to procedures in which mediation is offered during the administration of the assessment, usually whenever problems arise. As it was mentioned there are a number of other methods, such as Budoff's Learning Potential Measurement Approach, Guthke's Lerntest Approach, Carlson and Wiedl's Testing-the-Limits Approach, and Brown's Graduated Prompt Approach. Here we just mention their names and do not go into the details of their methodology, interested reader can refer to Poehner (2008) for more information.

Apart from these differences, it is the mediation phase of dynamic assessment which requires much care. This phase of the assessment is based on the Lev Vygotsky's concept of a ZPD. It is believe that interaction is a necessary precondition to learning (Nyikos and Hashimoto, 1997). The mediation stage reflects Vygotsky's ideas about instruction within the zone of proximal development (Vygotsky, as cited in Dixon-Krauss, 1996). It guides the teacher in making instructional decisions by analyzing the student, the text, and the type and amount of mediation she needs to provide. Thus, one may ask the question that how all these things can be attained in an academic setting. The answer is that all these things are achieved through *cognitive apprenticeship,* which is the topic of next section.

4. Cognitive Apprenticeship

According to Brown, Collins, & Duguid (as cited in Nyikos and Hashimoto, 1997) "Cognitive apprenticeship (CA) is the most frequently applied model of constructivist approaches" (p. 64). It strives to place teaching and learning practices within a rich and real-life context that is meaningful to students. Brill, Kim, and Galloway (2001) assert that

cognitive apprenticeship practices are educational approaches that strive, first and foremost, to place teaching and learning practices within a rich and varied context that is meaningful and authentic to students. It has certain characteristics such as modeling, scaffolding or coaching, and reflection.

Modeling starts from learning simple tasks, such as riding a bike which involves more physical skills to complex tasks that require cognitive processes. In a typical modeling strategy both teachers and students are serving cognitive role. The models should put their thoughts and reasons into words while explaining and demonstrating certain actions, because students cannot otherwise monitor the thinking process (Shunk, 2000, as cited in Brill, Kim, and Galloway, 2001). After all**,** the aim of modeling in cognitive apprenticeship is to show how a process unfolds. Scaffolding involves a teacher (or a more knowledgeable other) providing some type of assistance, help or support to a learner to facilitate attainment of a goal. In fact, scaffolding can be considered only one form of coaching (Brill, Kim, and Galloway (2001).

The cognitive apprenticeship model can also be implemented as reflection. The goal of reflection is that students have enough opportunities to look back and analyze their individual and group performance in order to have a more chance for understanding and improvement. Like other components of cognitive apprenticeship, reflection can be encouraged in students in a variety of ways. For example, a mentor can pose "why" questions. Cognitive apprenticeship with its emphasis on situated learning encourages authentic activity and assessment (Collins, Brown, & Newman, 1989). Because Cognitive apprenticeship provides students with authentic tasks it encourages them (Collins, 1991).

B. Empirical Studies on L2 Dynamic Assessment

In their comprehensive review of application of DA to educational settings Haywood and Lidz (2007) assert that "Campion and Brown (1987) have been pioneers in their attempts to assess specific academic domains in the framework of DA" (p. 77). Lantolf and Thorne (2006) believe that the entire body of research in this new avenue of research includes only few studies that focus on L2 learners or bilinguals. They begin their review with the work of Pena and Gillman (2000) (as cited in Lantolf and Thorne, 2006) who investigated the children's reasoning through DA. The second study which they refer to is that of Anton (2003) which uses DA as a placement procedure. Participants were asked to construct orally a past-tense narrative after watching a short video clip. This time the learners received no feedback or mediation. They were then shown a second clip and asked to repeat the task, but this time with the help of a mediator who offered suggestions, posed questions, made corrections and helped them think through decisions making. After approximately six weeks of instruction, the participants were re-administered the original independent and mediated narration tasks in order to cheek their progress.

Poehner (2008) also conducted a series of extensive DA case studies examining oral proficiency among advanced undergraduate learners of French. Another study is that of Kouzlin and Grab (2002) which is about the EFL reading comprehension of adult at-risk immigrants. The results of their study indicate that the procedure is both feasible and effective in obtaining information on students' learning potential. It is confirmed that students with a similar performance level demonstrate different, and in some cases dramatically different ability to learn and use new text comprehension strategies. One interesting aspect of their work is the manner in which they report the outcomes of the DA procedure. Rather than generating a qualitative report of each learner's performance for all stages of the study, they presented the learners' abilities with a single score which they themselves called Learning Potential Score (LPS) which is the difference between the learner's pretest and posttest scores.

Of other examples of the direct application of DA to the domain of language we can refer to the works of Roseberry and Collin (1991) and Jacobs (2001). The results of former study indicated that addition of intervention was effective. The results of the latter study also showed that inclusion of a dynamic component to preschool program developed the knowledge of preschool children. Bendar and Kletzian (1990) applied a pretest-intervention-posttest format to 29 students from grade 9 to 12 and they saw development in their reading.

Ableeva (2007) using a DA procedure in assessing listening comprehension of university level L2 learners of French uncovered the source of comprehension problems. He found that in one case student shifted to a single lexical item and in another one to cultural knowledge. This revealed that learners' abilities were more developed than one would have surmised from unmediated performance.

III. METHOD

A. Participants

There were 30 male EFL participants (fifteen in each class) who were selected based on the available sampling procedure. According to the final exams of the previous semesters of the language institution these students were supposed to enjoy the same proficiency level, and now they were at 12th semester of New Interchange course. To double-check the homogeneity of subjects in terms of reading comprehension skill the reading part of Nelson English Language Proficiency Test was administered. To compare the gained mean scores of the two groups on the test a t-test was run to ensure that control and experimental group are homogenous. The results of the t-test are presented in table 3.1.

Table 3. 1. Independent t-test (group copmarison)

| | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
| | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
								Lower	Upper
Equal variances assumed	1.554	.219	-.561	28	.577	-2.320	4.132	-10.63	5.989
Equal variances not assumed			-.561	26.166	.577	-2.320	4.132	-10.64	5.997

As can be seen, the t-observed value, i. e. -.56 and the critical t-value at 28 degree of freedom is 2.084. In statistical term, the amount of t-observed, i.e. -56 at 28 degree of freedom is much lower than the critical F value, i. e. 2.048. Thus, when the t-observed is smaller than the critical-t value, it can be concluded that there is no significant difference between the means of the two groups and we can be on the safe side to claim the homogeneity of groups prior to the mediation. One of the two groups was randomly assigned to the control group and the other to the experimental group.

B. Instruments

In this study five instruments were used. The first one was the reading section of Nelson English Language Proficiency Test which was adopted from a Thesis by Avaz zade (2005, pp. 55). It was used to measure reading comprehension of students and ensure their homogeneity. Overall this test consists of 30 items. Four other reading comprehension tests from TOEFL Reading Flash (2005) were also used. According to Broukal, the author of the book, these passages are easier than the passages on the TOEF, so they can be used in classroom for students who need more work on reading. Haywood and Lidz (2007) also assert that the content of pre and post test should be just beyond the individuals' zone of actual development so that requires intervention or mediation. One text was used as a standard pretest, the other as the dynamic post test for determining the zone of possible development and the other two tests were used during the intervention phase as the material of practice.

C. Design

First it should be pointed out that this small-scale classroom research was implemented to improve the learning potential of students. But due to the fact that we did not have a pool of subjects we had to adopt especial design. As Best and Kahn (1989; 116) maintain "randomization involves pure chance of selection and assignment of subjects to experimental and control groups for a limited supply of available subjects". In other words, randomization occurs at two levels of selection and assignment. We did not have the first type of randomization and the only option left for us was to randomly assign them to experimental and control groups.

The design of the study was based on a *Sandwich model* which was introduced by Sternberg and Grigorenko (as cited in Poehner, 2008). According to Poehner (2008) in this approach to DA, a mediation phase is sandwiched between pretest and posttest that are administered in a non-dynamic manner. The performance on the post test can then be compared to that of the pretest in order to decide how much improvement has been made. Sternberg and Grigorenko also point out that the mediation procedure can be administered in either an individual or group setting. So, based on these theoretical assumptions the author implemented these procedures a detailed description of which is presented in the next part.

D. Procedure

In this study we selected two homogenous EFL classes but as was said to ensure their homogeneity a Nelson reading test was administered. Then, they were assigned to control and experimental groups. During four successive sessions 30 minutes of class time was allocated to mediation and discussion about the results of their exams. During the mediation phase the researcher followed an interactionist method which is based on cooperative dialoging. In this approach, assistance emerges from the interaction between the mediator and the learner. As the methodology of DA required us we have to abandon the traditional conceptualizations of the examiner/examinee roles in favor of a relaxed, stress-free, and dynamic relationship in which both are working toward the ultimate success of the student. In this joint enterprise the teacher (and in this case mediator himself) provided guidance. The fact is that in this method teacher cannot be predesigned because it is not possible to know in advance what feedback and support would be required by the student. The mediator provided help first for the failing individual with the necessary information about the problem and the possible ways of solving it. Then, once the individual has developed the ability to tackle the problem (through mediation work with the teacher), their ability to tackle similar tasks were assessed.

In fact the mediator did not follow any specific procedure for instruction because as Poenher and Lantolf (2005) stated in *interactionist* DA "instruction is not organized according to a net sequence" (p.10). They also assert that tasks are presented according to teacher's assumptions and the teacher should be ready to provide appropriate mediation at the appropriate time. It includes presenting the problem to the students and let them think through the process. During mediation it was tried to help the experimental group so that they could reach their full potential or what is known as zone of proximal development. A description of each stages of mediation is provided.

Phase 1: In the pre-test stage the students were given a reading comprehension test and their papers were taken home and corrected. After correction they were scored and it was tried to provide necessary information via marginal comments or other possible ways. Then, the papers were taken back to the class and distributed among students to start the second phase.

Phase2: This stage of the study was the heart of mediation and the most burdensome part for the researcher. The main content of the second part or the mediation phase was based on the performance of the students in the pre-test. During the mediation it was tried to provide enough feedback and discuss a wide range of topics that could be helpful in text comprehension. The teacher assisted the students as they were trying to agree on the correct answer of the questions that they missed. In fact, there was a kind of discussion. The discussion was about the text structure, cohesion devices, coherence, reading strategies and textual metadiscourse devices.

The mediator also helped students to understand the requirement of text comprehension, offer feedback, let the students verbally report the answering strategy, explain the reason, and examine the strategies used. For examples, they questioned reasons for choosing choice A rather than B, C, or D, and the teacher explained and discussed how a correct answer was obtained. During this stage several questions were raised by students. The dialogic nature of mediation lead to the development of insight of the subject's understanding of the structure of text, text comprehension, and better test taking strategies. Following is a list of the most frequently applied strategies for giving feedback on their pretest

- Providing marginal comments on the pre-test paper and discussing them with the students
- Asking marginal questions
- Providing tentative hints
- Asking peers for help

The pretests were given back to the students. In additions to collaborative reviewing the above-mentioned factors, in an attempt to increase their knowledge of text structure, the researcher asked the students to

- Highlight or underline modality markers which are used to comment on the truth-value of the ideas (e. g. it seems that)
- Highlight or underline sequencers and topicalizer (e. g. first, the next problem is)
- code glasses which help readers grasp the meaning of words (e. g. Assiduous which means "hard-working").

Phase3: in this part of the study the fourth reading comprehension was administered as the dynamic post test. The procedure of scoring was the same as pre test. That is, students were given score based on an interval scale, and false answer did not have negative score. The procedure we followed for calculating the end result of our study was the same as the method Kuzlin and Grab (2002) used in their study. We also reported the outcomes of the DA procedure in terms of Learning Potential Score. Rather than generating a qualitative report of each learner's performance for all stages of the study, we presented the learners' abilities with a single score which is called Learning Potential Score (LPS is the difference between the learner's pretest and posttest scores. The results are presented in more details in next part

IV. RESULTS

The main purpose of this study was to find the effect of dynamic assessment procedure on reading comprehension performance of EFL students. Assuming that DA would have no significant effect on the reading ability of students, the following research question was proposed. The following is a recitation of the research question:

Is there a significant difference between reading comprehension performance of dynamically-assessed students and non-dynamically-assessed ones?

Based on this question the following null hypothesis was proposed:

There is no significant difference between reading comprehension performance of dynamically-assessed students and non-dynamically-assessed ones?

In order to probe the null hypothesis stating that there is no significant difference between the EFL students' performance on non-dynamic and dynamic reading assessment the two groups first took a standard or a non-dynamic reading test. After reviewing the results of their test and giving additional instruction during mediation phase, a dynamic post-test reading comprehension was conducted. It should be pointed out that our control group also took the same tests but instead of instruction in the mediation phase a placebo was introduce to them. In order to compare the mean scores of the control and experimental groups on the pretest and posttest, the gained mean score for each group was separately computed by subtracting the posttest mean score from the pretest mean score. An independent t-test was run to compare the difference in the mean scores of the two groups. Table 4.2. shows the results.

TABLE 4. 1.
DESCRIPTIVE STATISTICS

	GROUP	N	Mean	Std. Deviation
PRETEST	1.00	15	13.5333	2.1252
	2.00	15	14.0333	2.0482
POSTTEST	1.00	15	14.1667	1.3844
	2.00	15	17.3333	1.7995

TABLE 4. 2.
INDEPENDENT T-TEST

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Equal Variance assumed	3.763	.058	4.011	28	.000	2.17000	.24081	1.6858	2.65419
Equal Variances not assumed			3.011	43.217	.000	2.17000	.24081	1.6844	2.65558

As can be seen, the t-observed value is 4.01, and the critical t-value at 28 degree of freedom is 2.084. In statistical term, this amount of t-observed, i.e. 4.01 at 28 degree of freedom is bigger than the critical F value of 2.048. Thus, when the t-observed is bigger than the critical-t value, it can be safely concluded that there is a significant difference between the means of the two groups. Thus, the null hypothesis can be safely rejected and we can claim the effectiveness of instruction during the mediation phase.

Effect size

The information gained so far give us only an indication of whether the difference between groups is 'statistically significant' (i.e. not likely to have occurred by chance) or not. But, there is more to research than just obtaining statistical significance. One way that you can assess the importance of your finding is to calculate the 'effect size' (also known as 'strength of association'). Effect size statistics provide an indication of the magnitude of the differences between groups (not just whether the difference could have occurred by chance). There are a number of different effect size statistics, the most commonly used being eta squared. Eta squared can range from 0 to 1 and represents the proportion of variance in the dependent variable that is explained by the independent (group) variable. Calculating the eta squared value (effect size) for this research the obtained value was .36. The guidelines for interpreting obtained value are as following: .01=small effect, .06=moderate effect, .14=large effect. Based on these guidelines for interpreting effect size .36 shows a large effect size.

V. INTERPRETATION OF RESULTS AND DISCUSSION

The aim of this paper was to explore the feasibility and practicality of development and implementation of DA to reading comprehension of EFL students in the context of Iran. A comparison of mean score of experimental group displayed that the mean score of experimental group had an increase of 2.3 score in posttest while the increase for the control group was only 1.1. It reveals that the students in the experimental group performed better in comparison with the students in the control group. The finding is consistent with the finding of Kuzlin and Grab (2002), who provided evidence for the positive effects of DA on performance of students. Based on the results of the study it can be discussed that DA is more than just a formative assessment. As we have seen in our brief analysis of the data of teacher interact with students we saw the aim was to promote development. Thus, one possible explanation for the positive effect of DA on reading comprehension is that it is more than just a sheer form of assessment. DA is a pedagogical approach which is supported by theories of mind and development. It is an approach which stresses the inseparability of assessment and instruction. Adding DA to the testing setting reduces the stress, gives learners extra confidence and they feel that there is someone who cares about them when they get stuck.

As all of us, either as student or as teacher have experienced situations in which we knew something but because of some destructive factors such as stress or test anxiety our mind went blank. Having this in mind, if we can create through DA a stress-free situation in which students are assured that there is someone who will care about them if they get stuck, their performance will increase dramatically. This issue is blatantly important for Iranian students who often lack strategies for coping with reading task and easily fall behind in completing the demanding task of reading comprehension. In this regard, teachers' behavior plays an important role in the real performance of students. Through their interaction and communication they will facilitate and assist forward students in selection of the right choice.

VI. CONCLUSION

This paper aimed at determining the effect of dynamic assessment of reading comprehension. Although in this research we saw a change in the performance of our experimental group, one should keep it in mind that mediation phase during which there is a dialogic interaction or interplay between mediator and learner may not result in sudden change in performance. Moreover, DA is highly influenced by its reliance on the meditational skills of the examiner. Put together, we should remember that even if we see no dramatic change it should not be interpreted as an indication of the lack of development, instead it impacts learner development. In addition to the statistical findings of this research there are a number of advantages in implementing DA that we can say that the end justify the means. Of major advantages of DA which justify its implantation is its fairness. It should be pointed out that DA is an integral part of the assessment but not its entirety, because no one approach can provide adequate answer to all questions.

In this study we only worked with 30 male students. Future research can investigate the same research project with both male and female subject. In this study we used the sandwich model of mediation. Other researcher can use other models of mediation. Research on DA has mostly focused on expert-novice relationship, but future research can work on the peer-peer interactions as possible sources of mediation. And finally, here we worked only on one skill known as reading, while further research can open new avenues of research into other areas such as listening, writing, vocabulary acquisition and so forth.

REFERENCES

[1] Ableeva R. (2007). Assessing Listening for Development. In R. Alanen and S. Poyhonen (eds.) *Language in Action. Vygotsky and Leontievian legacy today*. Cambridge Scholars Publishing, pp. 352-379.

[2] Antoń, M. (2003). Dynamic assessment of advanced foreign language learners. Paper presented at the American Association of Applied Linguistics, Washington, DC, March, 2003.

[3] Avaz zade, M. (2005). The relationship Between the Use of Metacognitive Strategies by Iranian Pre-University Male Students and their Reading Comprehension Scores. Unpublished Thesis.

[4] Bendar. M. R., and Kletzien, S. B. (1990). Dynamic assessment for reading: A validation. Paper presented at the National Reading Conference, Miami, FL.

[5] Best, J., & Kahn, J. V. (1989). Research in education. (Sixth Edition). New Jersey: Prentice Hall.

[6] Brill, J., Kim, B., and C., Galloway (2001). Cognitive Apprenticeship From Emerging Perspectives on Learning, Teaching and Technology. Retrieved July 23, 2009, from http://projects.coe.uga.edu/epltt/index.php?title=Cognitive_Apprenticeship.

[7] Broukal, M. (2002). Toefl Reading flash: Essential Practice for High Reading Comprehension Scores. Peterson's Thomas learning.

[8] Campione, J. C. & Brown, A. L. (1990). "Guided learning and transfer: Implications for approaches to assessment" in Diagnostic Monitoring of Skill & Knowledge Acquisition, Ed. Frederiksen, et al., Lawrence Erlbaum Assoc., Mahwah, NJ.

[9] Cioffi, G. and Carney, J. (1983). Dynamic assessment of reading disabilities. *The Reading teacher,* 36, 764-768.

[10] Collins, A. (1991). Cognitive apprenticeship and instructional technology. In L. Idol & B.F. Jones (Eds.), *Educational values and cognitive instruction: Implication for reform* (pp. 121-138). Hillsdale, NJ: Lawrence Erlbaum Associates.

[11] Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum Associates.

[12] Derry, S. J. (1999). A Fish called peer learning: Searching for common themes. In A. M. O'Donnell & A. King (Eds.).

[13] Dixon-Krauss, L. (1996). A Mediation Model for Dynamic Literacy Instruction. *Journal of Russian and Eastern European Psychology*, *1*, 78-85.

[14] Haywood, H. C., and C. S. Lidz. (2007). Dynamic Assessment in Practice: Clinical and Educational Applications. Cambridge. Cambridge University Press.

[15] Herman, L. J. And S. A. Zuniga. (2005). Dynamic Assessment. Retrieved August 19, 2009, from http://www.answers.com/topic/dynamic-assessment.

[16] Jacobes, E. L. (2001). The effect of adding dynamic assessment component to a computerized preschool language screening test. *Communication Disorder Quarterly, 22(4), 217-226.*

[17] Kozulin & Garb, (2002). Dynamic assessment of EFL Text Comprehension of At-Risk Students. *School Psychology International*, Vol 23, I, pp 112-127.

[18] Lantolf, J.P. & M.E. Poehner. (2004). Dynamic Assessment: Bringing the Past into the Future. *Journal of Applied Linguistics 1*: 49-74.

[19] Lantolf, J.P. and S.L. Thorne. (2006). The Sociocultural Theory and the Genesis of Second Language Development. Oxford: OUP.

[20] Lidz, C.S. & Gindis, B. (2003). Dynamic assessment of the evolving cognitive functions in children with typical and atypical development. In A. Kozulin, V. Ageyev, S. Miller, & B. Gindis (Eds.). *Vygotsky's theory of education in cultural context* [pp. 99-116]. New York: Cambridge University Press.

[21] Nyikos, M., and Hashimoto, R. (1997). Constructivist theory applied to collaborative learning in teacher education: in search of ZPD. *The Modern Language Journal,* 81, iv.

[22] Poenher, M, E. (2008). Dynamic Assessment: A Vygotskian Approach to Understanding and promoting L2 development. USA. Springer Science.

[23] Poenher, M, E. and Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research* 9, 3, pp 233-265.

[24] Roseberry, C. A., and Connel, P. G. (1991). The use of an invented rule in the differentiation of normal and language-impaired Spanish-speaking children. *Journal of Speech and Hearing Research, 34, 596-603.*

[25] Sternberg, R.J. and E.L. Grigorenko. (2002). Dynamic Testing. The Nature and Measurement of Learning Potential. Cambridge: Cambridge University Press.

[26] Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. M. Cole, V. John-Steiner, S. Scribner, E. Souberman (Eds.). Cambridge, MA: Harvard University Press.

[27] Williams, M. & Burden, R. (1997). Psychology for language teachers: a social constructivist approach. Cambridge: Cambridge University Press.

[28] Woolfolk, Anita. (2004). Educational Psychology. (9th ed). Boston: Allyn and Bacon.

Mehdi Mardani was born in Iran, has received his BA in Literature, MA in tefl at Tehran Tarbyat Moalem University (TMU), and now is a Ph.D. candidate majoring tefl at the University of Isfahn, Iran.

He has been teaching English as a foreign language at different levels such as language institutes and different Universities including University of Isfahan, Medical University of Isfahan, and Petroleum University of Ahvaz. He has published a paper in issue 30, April, 2010 of the Journal of Language, Society and Culture. He has also got the acceptance of two other papers in the international journal of Iranian EFL which are in the process of publication. His main areas of interest are language testing, ESP testing, and discourse analysis.

Mansoor Tavakoli was born in Iran, got his Ph. D. in applied linguistics from University of Tehran and now is the assistant professor of University of Isfahan. He has been teaching TEFL courses at the University of Isfahan for 20 years. His research interests are second language assessment and L2 acquisition. He has published several articles in this respect.