



## Dynamic assessment versus static assessment: A study of reading comprehension ability in Iranian EFL learners

Behzad Nazari <sup>a \*</sup>, Saeed Mansouri <sup>b</sup>

<sup>a</sup> Islamic Azad University, South Tehran Branch, Tehran, Iran

<sup>b</sup> Islamic Azad University, Tonekabon Branch, Tonekabon, Iran

### APA Citation:

Nazari, B., & Mansouri, S. (2014). Dynamic assessment versus static assessment: A study of reading comprehension ability in Iranian EFL learners. *Journal of Language and Linguistic Studies*, 10(2), 134-156.

### Abstract

This study extends traditional or static assessment of reading comprehension in foreign language contexts and applies dynamic assessment (DA) to the development of learners' reading ability. To homogenize the research population (N= 250), an Oxford Placement Test (OPT) was administered. On the basis of the test results, the population was sorted into three groups of reading-low, reading-mid, and reading-high students. The participants of this study were a sample of 30 participants with the lowest level of reading comprehension proficiency randomly assigned into two groups of control and experimental. This study was conducted in the Nosrat Institute, located in Kermanshah Province in 2012. In order to determine their current zone or level of reading proficiency, a TOEFL reading comprehension pre-test was administered to both groups. Afterwards, DA training was applied to the EG in 9 successive 80-minute sessions on reading comprehension, Unlike EG, static assessment was applied to the CG. In the end, another TOEFL reading comprehension post-test was administered to the research groups to measure their reading comprehension performance level after their treatment. The statistical data analysis revealed that DA was statistically more significant and effective for the low skilled readers than static assessment.

© 2014 JLLS and the Authors - Published by JLLS.

*Keywords:* Assessment; dynamic assessment; static assessment; zone of proximal development; reading comprehension ability

## 1. Introduction

This paper lies in an interdisciplinary field of applied linguistics that includes second language acquisition, language pedagogy and sociocultural theory, as proposed by Russian psychologist and educator Lev Semenovich Vygotsky. More specially, it investigates the pedagogical application of Dynamic Assessment (hereafter DA), a testing approach nurtured by Vygotsky's theory to reading comprehension.

Given the varied and often conflicting responsibilities teachers face daily, it is not surprising that assessment issues may prompt an exasperated, "Why do we assess anyway?" 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. Questioning the purpose of assessment may seem rhetorical since it has become as naturalized a part of everyday life as television and supermarkets. Nevertheless, assessment specialists are increasingly reflecting on the reasons behind specific assessment practices as well as the role of assessment in society. Traditionally, assessment is

\* Behzad Nazari. Tel.: +98-935-681-6340  
E-mail address: [behzad.nazari1362@yahoo.com](mailto:behzad.nazari1362@yahoo.com)

benignly described as an information-gathering activity (e.g., Bailey, 1996). For instance, McNamara (2004) explains that we assess in order to gain insights into learners' level of knowledge or ability. From this perspective, it is difficult to understand why educators, including second language (L2) teachers, often refer to assessment as "a necessary evil."

Changes in language teaching methods have always brought about changes in language testing and assessment. As expected such changes in language teaching orientation did not leave the assessment orientation untouched and along with teaching methods their corresponding testing or assessment systems also moved from product-oriented to process-oriented ones. But still as a result of focusing on wide-spread and commonly used product-oriented testing, teachers and assessors have no access to the needed information on performance of different learners because the only data provided for them by product-oriented testing is the final result of that test as a single score and nothing about the details of test task performance is in hand.

To solve this problem, DA in language learning derived from sociocultural theory (SCT) of Vygotsky and his idea on cognitive development offers new insights into assessment in the language classroom by revealing hidden aspects of individuals' abilities in answering each test item. While the results of traditional non-dynamic assessment (NDA) or Static Assessment (hereafter SA) can only show the already existent abilities of the student, the analysis of zone of proximal development (hereafter ZPD) makes it possible to evaluate the ability of the student to learn from the interaction with a teacher or a more competent peer and predict their possible future development. Vygotsky (1978) defined ZPD as the distance between a child's "actual developmental level as determined by independent problem solving" and the higher level revealed in "potential development as determined through problem solving under adult guidance or in collaboration with more able peers" (p. 86). Unaided performance on static measures tells us what has already been learned or accomplished, whereas the breadth of ZPD is thought to provide prospective indications of what can be learned. While studying the development of children's mental abilities, Vygotsky (1978) observed that what a child is able to do independently only displays the tip of iceberg, that is, a partial picture of child's full capability, because the child can often do more when just a bit of assistance, or mediation, is offered by someone else. According to Vygotsky, what the child is able to do automatically shows a view of the child's past development, but what the child is able to achieve with mediation, provides insights into the child's future development. The mediator facilitates learning, allowing the access and the unfolding cognitive functions that the subject has not yet mastered: "... what the child can do today with the help of an adult, it will do tomorrow without any help" (Vygotsky, 1979, as cited in María del Carmen Malbrán; Claudia M. Villar, 2002, p.2)

This study intends to investigate that whether making strategies awareness (mediation) at the particular level of examinees cognitive development can assist examinees to figure out and to apply those strategies that potentially exist in their cognitive structures as capabilities, but never explicitly emerged in their true abilities. To put it another way, this study investigates the feasibility of the development and implementation of the DA procedures in Teaching English as a Foreign Language (TEFL). The results of the study may illustrate whether dynamic procedure provides any information on students' learning over and beyond what is available from SA.

The present study in the domain of DA investigated reading comprehension performance of 30 Iranian EFL students in Nosrat Institute in Kermanshah whose age ranged from 16 to 20. They were high school students or college students attempting to learn English as a foreign language. There were thirty students enrolled in the course, and participated in the entire research process who were evenly divided in to two experimental and control groups. The major question is whether DA training affects the individual's reading comprehension performance.

## 2. Literature Review

The theoretical roots of Dynamic Assessment lie in Vygotsky's sociocultural theory of child development (Vygotsky, 1986) in which the role of the parent, teacher, sibling, or peer, in interacting with the individual child, is seen as fundamental to the formation and growth of cognitive skills, which are culturally mediated through these interactions.

One of the fundamental concepts of sociocultural theory, according to Lantolf (2000), is its claim that the human mind is mediated. Lantolf claims that Vygotsky finds a significant role for what he calls 'tools' in humans' understanding of the world and of themselves. According to him, Vygotsky advocates that humans do not act directly on the physical world without the intermediary of tools. Whether symbolic or signs, tools according to Vygotsky are artifacts created by humans under specific cultural (culture specific) and historical conditions, and as such they carry with them the characteristics of the culture in question. They are used as aids in solving problems that cannot be solved in the same way in their absence. In turn, they also exert an influence on the individuals in that they give rise to previously unknown activities and previously unknown ways of conceptualizing phenomena in the world. Therefore, they are subject to modification as they are passed from one generation to the next, and each generation reworks them in order to meet the needs and aspirations of its individuals and communities. Vygotsky advocates that the role of a psychologist should be to understand how human social and mental activity is organized through culturally constructed artifacts.

According to Vygotsky (1978 cited in Lantolf 2000), the sociocultural environment presents the child with a variety of tasks and demands, and engages the child in his world through the tools. In the early stages, Vygotsky claims that the child is completely dependent on other people, usually the parents, who initiate the child's actions by instructing him/her as to what to do, how to do it, as well as what not to do. Parents, as representatives of the culture and the conduit through which the culture passes into the child, actualize these instructions primarily through language. On the question of how do children then appropriate these cultural and social heritages, Vygotsky (1978 cited Wertsch 1985) states that the child acquires knowledge through contacts and interactions with people as the first step (interpsychological plane), then later assimilates and internalizes this knowledge adding his personal value to it (interpsychological plane). This transition from social to personal property according to Vygotsky is not a mere copy, but a transformation of what had been learnt through interaction, into personal values. Vygotsky claims that this is what also happens in schools. Students do not merely copy teachers capabilities; rather they transform what teachers offer them during the processes of appropriation.

Vygotsky (1978 cited in Lantolf 1994, 2002) argues that the field of psychology has deprived itself of crucial information to the understanding of complex aspects of human behavior by refusing to study consciousness. This refusal, according to him, has restricted the role of psychology to just the explanation of the most elementary connections between a living being and the world. Consciousness in his view distinguishes human behavior from other living beings and links the individual's knowledge to his/her behavior. It arises, functions and develops in the process of people's interaction with reality on the basis of their socio-historical practices. He insists that socially meaningful activity has to be considered as the explanatory principle for understanding consciousness and he rejects any attempt to decouple consciousness from behavior.

Lantolf et al. (1994) indicate that the latter understanding of consciousness in the field of teaching is embodied in the concept of metacognition, which, according to him, incorporates functions such as planning, voluntary attention, logical memory, problem solving and evaluation. Williams and Burden (1997) claim that sociocultural theory advocates that education should be concerned "not just with theories of instruction, but with learning to learn, developing skills and strategies to continue to learn, with making learning experiences meaningful and relevant to the individual, with developing and

growing as a whole person". They claim that the theory asserts that education can never be value-free; it must be underpinned by a set of beliefs about the kind of society that is being constructed and the kinds of explicit and implicit messages that will best convey those beliefs. These beliefs should be manifest also in the ways in which teachers interact with students.

Sociocultural theory has a holistic view about the act of learning. Williams & Burden (1997) claim that the theory opposes the idea of the discrete teaching of skills and argues that meaning should constitute the central aspects of any unit of study. Any unit of study should be presented in all its complexity rather than skills and knowledge presented in isolation. The theory emphasizes the importance of what the learner brings to any learning situation as an active meaning-maker and problem-solver. It acknowledges the dynamic nature of the interplay between teachers, learners and tasks and provides a view of learning as arising from interactions with others. According to Ellis (2000), sociocultural theory assumes that learning arises not through interaction but in interaction. Learners first succeed in performing a new task with the help of another person and then internalize this task so that they can perform it on their own. In this way, social interaction is advocated to mediate learning. According to Ellis, the theory goes further to say interactions that successfully mediate learning are those in which the learners scaffold the new tasks. However, one of the most important contributions of the theory is the distinction Vygotsky made between the child's actual and potential levels of development or what he calls Zone of Proximal Development (ZPD) defined as the distance between learners' existing developmental state and their potential development. Put another way, the ZPD describes tasks that a learner has not yet learned but is capable of learning with appropriate stimuli. The ZPD is an important facet of sociocultural theory because it describes tasks "that child cannot yet do alone but could do with the assistance of more competent peers or adults" (Karpov & Haywood, 1998).

Vygotsky claimed that instruction and assessment are only good when they promote development and stimulate a range of functions that are ripening within the ZPD (Vygotsky, 1978). Vygotsky's understanding of instruction and development of cognitive abilities served as the basis for creation of educational approaches that seek to target learners' ZPD. According to Vygotsky, the development of the child involves the appropriation of humans' cultural experience in collaboration with adults and includes two levels, i.e. actual level and potential level of development. The actual level presumes the child's independent problem solving and corresponds to zone of actual development. The potential level of development presupposes adult-child collaboration during problem-solving activities. These learning activities are intended to reveal the child's abilities that are in the process of maturation. The potential level is associated with the ZPD and is understood by Vygotsky as: the distance between the actual development level as determined by independent problem-solving and the level of potential development as determined by problem-solving under adult guidance or in collaboration with more capable peers (Vygotsky, 1978, p. 86).

Simply put, for Vygotsky, exploring what the child [the learner] can do independently, one explores the previous or actual level of child's [the learner's] development. But exploring what the child [the learner] is able to do with a more skilled other (e.g. parents, peers, teachers), one can determine the child's [the learner] potential development.

In his works, Vygotsky demonstrated that the ZPD of mentally retarded children is fairly small and for this reason, they have a limited capacity to appropriate assistance. However, this capacity may be quite significant among children with temporary development retardation. Within the testing framework, the ability to use adults' assistance as well as the ability to internalize this assistance and to transfer it to independent problem-solving is viewed as a positive diagnostic sign. This sign indicates that given child is not mentally retarded and that an appropriate instructional intervention can help the child overcome temporary mental retardation.

Lidz and Gindis (2003) state, “DA is an approach to understanding individual differences and their implications for instruction that embeds intervention within the assessment procedure. The focus of most dynamic assessment procedures is on the process rather than on the product of learning” (p. 99). In other words, in DA the mediator seeks to improve learner performance through modification of student activity. This interaction focuses on learner behavior and learner receptivity to mediation (Lidz, 1991).

Feuerstein, Rand and Hoffman (1979) suggested that DA differs from traditional standardized methods of psychological and psycho-educational assessment on several dimensions: Tzuriel summarizes them as follows: (Seng et al. 2003)

### *2.1. Experiments on Dynamic Assessment*

A statistical analysis of the pretest and posttest scores revealed that students did more than one standard deviation better on the posttest than they did on the pretest. This shows, according to Kozulin and Garb (2002), mediation was beneficial to students and that they were able to apply the strategies to which they were exposed in the mediation phase to novel situations. Moreover there was a negative correlation between the gain scores and pretest scores. In the opinion of the authors, this shows pretest scores do not reflect the students learning potential but rather their actual development.

The study revealed that pretest scores do not accurately explain a student’s ability to learn reading comprehension strategies. In fact, a closer examination of student scores reveals that students, who would have been classified at the same ability level according to a traditional placement test, instead have different developmental needs concerning text comprehension abilities.

Campione and Brown (1987) as cited in (Haywood and Lidz, 2007) have been pioneers in their attempts to assess specific academic domains in the framework of DA. Although primarily known for their highly structured approach to DA of mathematics learning, they, and primarily their student Palinscar, also designed a highly clinical appraisal of reading in the context of their reciprocal teaching model (Palinscar & Brown, 1984). Students are observed during reading comprehension groups regarding their abilities to summarize, formulate questions, clarify inconsistencies, and predict what will happen next. The teacher initially models these skills, increasingly releasing the role of group leader to the students. During this process, the teacher continues to model and provides feedback and prompts to enhance the students’ development of competence. The possibilities for transfer are maximized because the activities take place in the context to which they are to apply (Campione & Brown, 1985). There is no formal prescription for these interactions and no structure for recording of observations, although it would not be difficult to devise informal ratings or formats for anecdotal records.

Ruijsenaars and Oud (1987, in Hamers, Pennings, & Guthke, 1994) as cited in (Haywood and Lidz, 2007), also pioneers in this area, studied a long-term procedure in which they compared the effects of two reading programs with groups of kindergarten children. Intervention was provided by the classroom teacher during fifteen 10-minute sessions. The students were given pretests, as well as posttests that followed each of the lessons, a retention test 2 weeks following the completion of the sessions, and a reading speed test 5 months after completion of the sessions. The researchers found that 49% of the

variance in the final criterion measure could be accounted for by the posttests from the final learning session.

Tissink, Hamers, and Van Luit (1993) as cited in (Haywood and Lidz, 2007) applied their learning potential model to specific academic domains, including reading and spelling. Using a pretest–intervention–posttest format, these researchers sought to produce information that would predict later school achievement, profile strengths and weaknesses in children’s cognitive functioning, and provide information relevant to classroom instruction. Their domain-specific procedure relevant to both reading and spelling is the Auditory Analysis Test. The components of this procedure include memory for sentences, auditory segmentation of words in a sentence and syllables within a word, “objectivation,” to indicate which of two words is the longer and repeating a word leaving out one syllable, isolation of the first phoneme of words, and phonemic analysis segmenting words into their phonemes. Both pretest and posttest consist of 40 items, with 15 transfer items added to the posttest. The intervention offers a series of nonstandardized prompts (repeated presentation, revelation of item structure, provision of solution strategy, and modeling) that are offered as needed in response to learner errors. This procedure is appropriate for children in kindergarten and first grade, or any other children who are in the early stages of reading development. The authors provide data regarding construct, concurrent, and predictive validity; item analysis; and reliability. The domain specific tests were found to be better predictors of school achievement than the domain-general. Also, learning curves of the students differed per domain; that is, there did not seem to be a general trait of learning potential; rather, domain-specific responsiveness to intervention.

The Abbott, Reed, Abbott, and Berninger (1997) study of 16 second graders with severe reading and writing disabilities offers yet another approach within the broader DA model. These authors framed their study within the concept of “response to intervention” using growth curves as measures, with their tutorial intervention spanning a full program year. The tutorial targeted orthographic and phonological awareness, accuracy and fluency of word recognition and comprehension monitoring, handwriting fluency, spelling, and composing. The sessions began during the summer with 1-hour twice-weekly meetings over the course of 8 weeks; this continued through the school year with 1-hour once-weekly individual meetings over an average of 36.5 weeks. The lessons were standardized and scripted. This study showed gains beyond chance for most of the children on most of the measures but also showed their differential responses to the treatment. The results also generated individualized recommendations for their next school year.

## 2.2. *Statement of the problem*

Reading comprehension occupies a vital role in achieving success in all areas of life, in general, and academic life, in particular. Reading comprehension is a complex construct involving multi-level processes. In order to comprehend successfully, that is, to gain meaning from written text for a particular purpose, the reader must engage in various processes at the word-, sentence- and text level. The reader is required to identify a series of letters as a word, access the meaning of words and integrate individual

word meanings or sentence meanings into coherent sentence- and text-level representations (Best, Rowe, Ozuru and McNamara, 2005). Due to the intricacy of these prerequisites for effective comprehension is often the area where students with learning disabilities experience the most difficulty. Some of the problems for comprehending written texts originate from the traditional or Static Assessment (hereafter SA) in which the examiner is a neutral or "objective" participant who provides only standardized direction and does not give performance-contingent feedback. Indeed, the traditional assessment examiner is often explicitly discouraged from making any statements that may alter the independent achievement of the students.

Among the specific procedures currently used for reading placement are the informal reading inventory, basal placement tests, and standardized achievement tests. Each uses a static testing paradigm, a paradigm where a student's level of functioning is determined independently of any social interaction between the student and teacher. The teacher, following standardized procedure, administers these tests which presumably provide an objective measure of the student's reading ability. This test paradigm treats reading ability as a discrete, decontextualized trait. The students are allowed no support or aids throughout the testing period. The students cannot ask friends or the teacher how to pronounce a word or look at other resources to help determine the answer. These static tests, while useful for measuring what the students can read and understand without support, do not indicate the levels of reading materials that students can read with teacher support. Further, under the static paradigm, no information is provided regarding the amount or types of instruction needed to maintain adequate reading progress within the context of a reading lesson.

Bransford, Delcos, Vye, Burns and Haselbring (1987) described three reasons for moving away from traditional assessment or SA and moving toward Dynamic Assessment (hereafter DA). First, traditional assessment is only concerned with the products rather than processes of learning. Second, it fails to address each child's responsiveness to instruction because it is based on the premise that prior learning adequately predicts future performance. Third, it does not provide prescriptive information for designing potentially effective instruction. In other words, traditional assessment focuses on the product rather than the process of learning. It also emphasizes the outcomes rather than the strategies for learning and offers no information on the child's responsiveness to teaching, learner's future learning potential, or pedagogical needs. In line with these reasons, Jitendra and Kameenui (1993) argued that the failure of SA has prompted researchers to search for new assessment approaches designed to be more responsive to individual learner's potential strengths and weaknesses. The reason is traditional assessment is not designed to evaluate specific instructional strategies for remediation of learning deficits. It also does not recognize the learner's potential to succeed with adequate environmental support. To help alleviate this dissatisfaction with traditional assessment, many researchers have attempted to focus on DA to provide more information about the individual's learning ability. DA has emerged from both theoretical conceptions about human cognitive plasticity and practical needs to find novel diagnostic measures for children, unable to reveal their capacities in traditional assessment. On the other hand, Cioffi and Carney (1983) argued that SA procedures are best at evaluating the students' knowledge of skills, but insufficient for estimating the students' learning potential and provide little help for identifying the conditions under which the progress can be made. Finally, Lunt (1993) stated that the intention of traditional assessment or SA is to measure actual development, which is often subsequently confused with and used as measures of potential. DA procedures, on the other hand, involve a dynamic interactive exploration of the learner's learning and thinking processes, and aim to investigate a learner's strategies for learning and ways in which these strategies may be extended or enhanced. Since it offers individuals an opportunity to learn, DA has the potential to show important information about individual strategies and processes of learning. Therefore DA offers potentially useful suggestions about teaching. In sum, comparisons of SA versus DA could be described as in Table 1.

**Table 1.** Comparisons of static assessment versus dynamic assessment

Static Assessment	Dynamic Assessment
Examiner is an observer	Examiner is a participant
Examinee receives no mediation	Examinee receives mediation
Diagnosis $\neq$ Instruction	Diagnosis = Instruction
Focused on product	Focused on both process and product
Retrospective approach	Prospective approach
Decontextualized	Contextualized
Low transfer test $\neq$ Authentic task	High transfer test = Authentic task

DA procedures involve a dynamic interactive exploration of a learner's learning and thinking process and aim to investigate a learner's strategies for learning and ways in which these may be extended or enhanced. Since it offers individuals an opportunity to learn, DA has the potential to show important information about individual strategies and processes of learning and, therefore, to offer potentially useful suggestions about teaching.

While transferring DA techniques from purely cognitive functions to content learning (like language) certain factors must be born in mind. Unlike general cognitive functions associated with content-neutral logical reasoning that are to be fluid and amenable to change, the functions associated with reading are usually described as "crystallized" (Kozuilm and Garb, 2002) and resistant to short term changes. In the field of EFL reading comprehension, standard reading tests contain a large amount of materials (e.g., vocabulary and grammar for production), comprehension of which depends on students' previous knowledge rather than cognitive functioning (Kozuilm and Garb, 2002). Whereas DA tests of general cognition could be constructed by using the material of standard psychometric tests, DA in content areas requires the construction of special materials sensitive to both, more specific materials and more cognitive strategies use.

### 2.3. Purpose of the study

The purpose of this research is to study the Dynamic Assessment of reading comprehension and investigate whether this kind of assessment can have a positive effect on reading comprehension. Because of dissatisfaction with the use of IQ tests and discrepancy models over the past 10 to 15 years, researchers and educators have been investigating more efficient methods such as DA to improve learning ability.

Since reading is the basis of instruction in a social studies classroom, students must be able to comprehend the main ideas and themes of what they have read. Unfortunately, many students are unable to differentiate between what is important and what is not when they read primary or secondary source materials in class or for homework. Much of the literature states that DA treatment is often seen as useful tools to promote student reading ability.

### 2.4. Research questions and hypotheses of the study

The researchers investigated the following research questions, their hypotheses, and null hypotheses. The null hypotheses are those used for statistical analysis where significance is set at  $p \leq .05$  for this investigation.

1. Does Dynamic Assessment (DA) affect Iranian EFL students' reading comprehension performance?

*Null Hypothesis 1: Dynamic Assessment (DA) does not affect Iranian EFL students' reading comprehension performance.*

To examine the difference in the effects of DA on the research groups before and after the DA treatment, the following two questions, hypotheses, and null hypotheses were also developed:

2. Will the Experimental Group (EG) of the study show any progress from the pre-test to the post-test of the study?

*Null Hypothesis 2: The EG of the study will not show any progress from the pre-test to the post-test of the study.*

3. Will the control group of the study show any progress from the pre-test to the post-test of the study?

*Null Hypothesis 3: The Control Group (CG) of the study will not show any progress from the pre-test to the post-test of the study.*

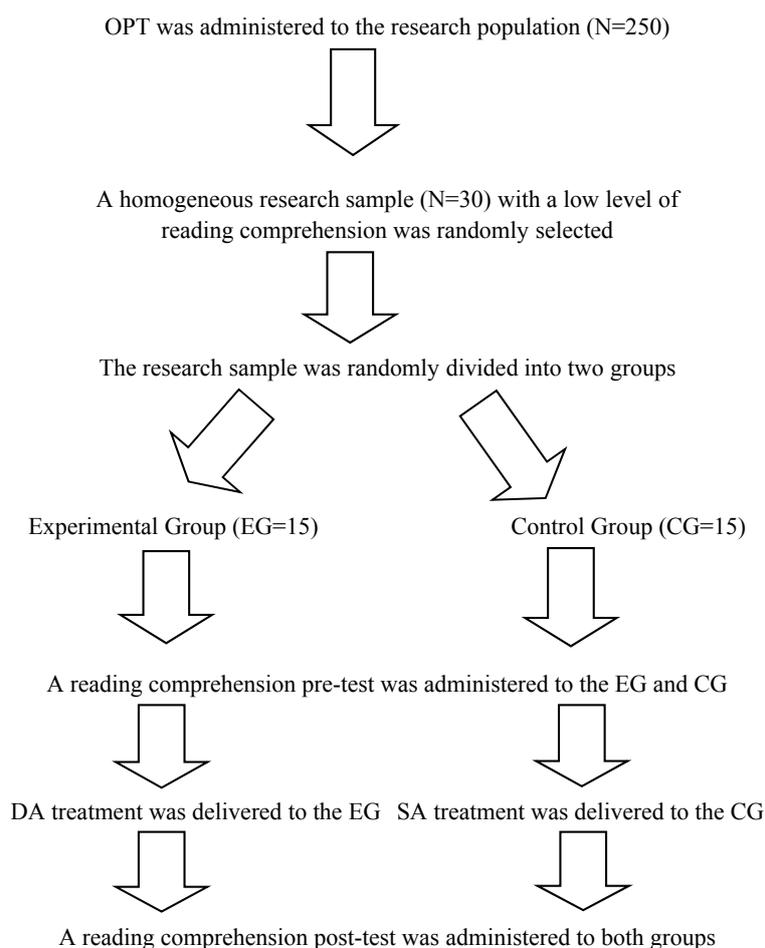
To answer these questions, a quasi-experimental research was carried out, including a pre-test (i.e. a TOEFL reading comprehension test), 9 successive 80-minute sessions training oriented towards DA originally designed for the EG and static assessment mainly intended for the CG, and a post-test i.e. a TOEFL reading comprehension test. The training on the use of DA in the EG and the use of SA in the CG provided knowledge of and practice on using GOs and Non-GOs respectively. The pre-test results helped determine the participants' current level of reading comprehension proficiency. The post-test results revealed the students' reading comprehension performance after the DA training. It should be added that the Oxford Placement Test aimed at homogenizing the research participants.

### 3. Method

The pre-test-intervention-post-test quasi-experimental design was used to investigate the effectiveness of using DA to enhance reading comprehension skill. This kind of design is one of the most frequently used designs in social studies research. It is structured as a pre-test-treatment-post-test experiment. The design of the research study includes four chronologically separate stages: OPT used to homogenize and select the research sample, the TOEFL reading comprehension pre-test intended to determine the current level of reading comprehension of the research participants; the training to use DA in the EG and SA in the CG; and the administration of the TOEFL reading comprehension post-test in order to determine the reading proficiency level of the EG and CG after DA and SA instructional intervention respectively. All the tests i.e. OPT, pre-test, training, and post-test administered in the classroom setting during regular class time. The participants took the OPT in first class session and the pre-test in the second session, and received DA training in the 9 subsequent, successive class sessions.

The post-test took place in the twelfth class meeting after the training, and it completed the treatment process. As a matter of fact, the study took 12 successive 80-minutes class meetings that is, the first class meeting was for the OPT; the second for the pre-test, from the third to the eleventh were for the training; and the last but not the least was for the post-test). The schedule for participants from the control group, on the other hand, was within the same number of class meetings, but received SA instruction. Because the researchers wanted to investigate the effect of DA on reading comprehension upon those students who have low-reading ability, students were required to have the same achievement level i.e., low-reading ability in reading comprehension. It should be mentioned that each class was instructed in the same curriculum for the same amount of time, and had similar number of participants.

DA was used in the EG, while the CG received instruction on the same book and content using SA. They were in the same learning communities but different classes in the Nosrat Institute, located in Kermanshah as mentioned beforehand. Figure 1 depicts the research design, demonstrating the pre-test, intervention, and post-test.



**Figure 1.** The Diagram of the Design of the Study

### 3.1. Participants

The research participants for the present study included the researchers as the mediators and two groups (i.e. the EG and CG) of students. A total of 30 students of English Nosrat Institute, located in Kermanshah province in Iran were randomly selected. Their age ranged from 16 to 20. It is worth mentioning that they did not remain constant throughout the data collection process that is, 4 participants (2 in the EG, and 2 in the CG) were absent on the tests, so they were excluded from the data analysis process. For the portion of the classroom instruction that involved DA in the EG and SA the CG, the researchers were the sole instructors of DA. All the participants were native speakers of Farsi and were chiefly from middle socio-economic background. Students' participation in the classroom activities and how prompting them to do so were the main researchers' concerns in the classroom. It should be mentioned that all the participants were male in gender.

As displayed in Figure 1, the number of the participants in each group was 15, but 2 students in the EG and CG did not take the research tests, and were therefore excluded from the study. Data on a student participant was valid only when s/he participated in all stages of the study. The ages of CG ranges from 15 to 21 but their ages mean is, to some extent, similar to EG. It is worth mentioning that, in the procedure section, just the participants' numbers were mentioned, that is, student 1 (S1), students 2 (S2) and so on.

### 3.2. Instruments

To collect quantitative data for the present study, three instruments were utilized: (1) a homogeneity test (i.e. the OPT); (2) a pre-test (i.e. a TOEFL reading proficiency test; and (3) a post-test (i.e. a TOEFL reading proficiency test). The OPT was used to homogenize the research sample. The results helped to sort the research population into three groups: reading-low (33%), reading-mid (33%), and reading-high (33%) groups. The sorting was based on percentile ranking. Because the researchers wanted to investigate the effect of DA on students who have low-reading ability, students in the group with the lowest level of reading proficiency was chosen as the research sample. Their current reading proficiency level was measured by the pre-test and the hypothesized development in their reading comprehension ability after the treatment was determined by the post-test.

#### 3.2.1. Oxford Placement Test

To homogenize the population, an Oxford Placement Test was utilized before the research intervention (DA training in the EG, and SA training in the CG). Some 30 EFL Iranian EFL learners with similarly low achievement levels in reading proficiency were selected. The reading comprehension ability before and after the DA-treatment in the EG and the SA-training in the CG was the major concern.

#### 3.2.2. Pre-test and post-test

In order to identify the testees' current level of reading comprehension, a reading comprehension test was administered as one of the primary sources of data for this investigation (See Appendix B). The 30 students with low-reading ability participated in the pre-test (15 in the EG, and 15 in the CG). It included twenty multiple choice items for two reading passages. These passages were selected from Longman Complete Course for the TOEFL TEST (Deborah Philips, 2005). The passages have different topics and almost of the same length and number of test items. The purpose of giving different reading subjects was to avoid topic bias and topic familiarity. Time allotment for the pre-test was forty-five minutes. In assessment, one point was awarded for a correct item.

Results of the post treatment test were compared with those of the pre-test in order to make inferences on the effectiveness of the DA training in the EG and the SA training in the CG through the change in the students' reading comprehension performances. The post-test in the EG was taken by thirteen out of the fifteen students who had gone through the pre-test and the DA training. Moreover, the post-test in the CG was taken by thirteen out of the fifteen students who had taken the pre-test and received the SA treatment (See Appendix C). It should be mentioned that the researchers used a pilot study in order to determine the reliability and the validity of pre-test and post-test. Five students were randomly chosen and after the test, the researchers used KR-21 method to determine the reliability of the test.

### 3.3. Procedure

To homogenize the research population (N= 250), an Oxford Placement Test (OPT) was administered. On the basis of the test results, the population was sorted into three groups of reading-low, reading-mid, and reading-high students. As many as 30 participants with the lowest level of reading comprehension proficiency was randomly selected and assigned to an Experimental Group (N=15) and a Control Group (N=15). This study was conducted in the Nosrat Institute, located in Kermanshah Province in 2012. In order to determine their current zone or level of reading proficiency, a TOEFL reading comprehension pre-test was administered to both groups. Afterwards, DA training was applied to the EG in 9 successive 80-minute sessions on reading comprehension, Unlike EG, static assessment was applied to the CG. In the end, another TOEFL reading comprehension post-test was

administered to the research groups to measure their reading comprehension performance level after their treatment.

The goal of this section was to focus on the mediational move introduced by the researchers (or mediators) during the study. The amount of mediation provided to each participant depended on the number of incorrect answers to each question. The participants received mediation on the questions they answered incorrectly. Therefore, the more incorrect answers a participant gave to the question, the more mediation he received. In addition, based on Aljaafreh and Lantolf (1994), intervention should be provided in gradual progression. The purpose is to estimate the minimum level of guidance required by the students to successfully perform a given question. Therefore, guidance normally starts at a highly strategic (implicit) level, and progressively becomes more specific, more concrete, until the appropriate level is reached. In order to make the dynamic assessment more systematic, in this study the researchers designed a series of mediation which began with the most implicit hints to the most explicit ones.

At the outset of the each session, the students were asked to pose question whenever they encounter any problem with reading comprehension. The mediations used in each session included hints, prompts, questions, and explanations. The researchers also introduced different kinds of strategies that could be used to answer a reading comprehension question. In the five first sessions, for example, the texts were followed by detail questions; the researchers clarified the strategies for answering these kinds of questions correctly. It should be mentioned that when the explanations were difficult for students, they were explained in their native language, in this case Persian.

The researchers also helped students to understand the requirement of text comprehension, offered feedback, let the students verbally report the answering strategy, explain the reason, and examine the strategies that were introduced. Moreover, in order to make interaction possible during mediation, the researchers allowed the students to use their native language.

In order to work on vocabulary, the researchers or rather the mediators introduced new techniques to the learners in order to learn unknown vocabularies such as associating a certain word with a routine activity, visualizing and making mnemonics for words, paying attention to the second meaning of words, thinking logically about the relation between words and sentences, making an educated guess through context clues, making flash cards and using them here, there, and everywhere, posing a bunch of flash cards around yourselves (peripheral learning), and making a notebook and learning how to use a dictionary.

### *3.4. Data analysis*

To answer the research questions an Independent Samples T-Test, Pearson Product Moment Coefficient Correlation (PMMCC), paired sample t-test, and independent sample t-test were used. An independent one-sample t-test was used to test whether the average of a sample differ significantly from a population mean, a specified value  $\mu_0$ . In order to investigate whether there was any difference between the pre-test and post-test scores of control group on one hand and pre-test and post-test scores of experimental group on the other hand, paired sample t-test and estimate of correlation between each group was used. For estimating the difference between the group receiving treatment, experimental group, and the group not receiving treatment, control group, Independent sample T-test was used to see whether the difference was meaningful or not.

## **4. Results**

This section is oriented towards the descriptively and inferentially statistical analysis of the quantitative data and findings gathered through two major instruments of the present research study (that is, the pre-test and post-test in the two independent groups of the study). The analysis was performed in

the light of two different but related branches of statistics: descriptive and inferential statistics. Each will be presented and discussed.

#### 4.1. Descriptive Statistics

Before examining the descriptive results obtained from the descriptive statistics, it should be borne in mind that in any empirical study a set of statistical assumptions should be satisfied in order to interpret the statistical results gathered from the data collection process. The following four SPSS outputs represent one of these assumptions showing the statistical analysis of the scores gained on the pre-test and post-test for the research groups (See Figure 4.1., and Figure 4.2).

**Table 2.** One-Sample Statistics for the pre-test and post-test scores of experimental group

	N	Mean	Std. Deviation	Std. Error Mean
Pre-test EX	15	8.4667	2.13363	.55090
Post-test EX	13	23.1538	23.1538	2.74527

Table 2 pertains to descriptive statistics. It is the first output of the T-Test. It represents the results of the descriptive analysis showing how the participants (N=13) performed on their pre-test and post-test as well as the dispersion of the scores on the two tests. Their performance is manifested by the mean scores. The first row represents the normal parameters and descriptive statistics including the mean, standard deviation, and standard error mean. The mean scores of the pre-test and the post-test are M= 8.46 (N= 15, SD= 2.133) and M= 23.1538 (N=13, SD= 23.15) respectively. The difference between scores on the pre-test and post-test in terms of the two mean scores, the standard deviation values, and the standard error mean scores indicates that the group performed differently before and after the training. Therefore, it is of high importance to study these differences by inferential statistics displayed in Tables below to answer the main research question of the present study.

**Table 3.** One-Sample Statistics for the pre-test and post-test scores of control group

	N	Mean	Std. Deviation	Std. Error Mean
Pre-test CO	15	9.0000	2.42015	.62488
Post-test CO	13	10.6154	1.93815	.53755

Table 3 is also related to descriptive statistics. As shown in it, the mean scores of the pre-test and the post-test in the CG are M=9.00 (N=15, SD=2.42), and M=10.61 (N=13, SD=1.93) respectively. The difference between the two mean scores, the standard deviation values, and the standard error mean scores on the pre-test and post-test shows that this group of participants, like the EG members, performed differently before and after static assessment, but how much change they underwent and whether the change was statistically significant in comparison with the EG will be studied in the following tables of Inferential statistics.

#### 4.2. Inferential Statistics

Having calculated the descriptive statistics based on the participants' scores on the pre-test and post-test, the researchers conducted some other data analysis statistical methods including the Paired Samples T-Test, the Independent One Sample T-Test, the Pearson Product Moment Correlation Coefficient (PPMCC) Test. The results of the each method will be presented and described.

Table 4. Paired Samples Test

Paired Differences									
95% Confidence Interval of the Difference									
	Mean	Std. Deviator	Std. Error Mean	Lower	Upper	t	df	Sig.(2-tailed)	
Pair1 pre test EX- post test Ex	14.53846	8.21194	2.27758	-19.50088	-9.57604	6.383	12	.000	

Table 4 represents the results of the T-Test for the EG of the study. As it is clear, it shows the comparison made between the pre-test and the post-test mean scores for the EG. The purpose of this comparison is to find out whether the participants in the EG made changes in their reading comprehension performance after Dynamic Assessment, and if so, how much change they underwent after the DA. To the right of the Paired Differences, the T (6.38), degrees of freedom (12), and significance (.000) are represented. Because the p-value is below 0.05, the H0 (there is no significant difference between the means of the two variables) is rejected. Here, it is seen that the significance value is statistically significant. There is a great difference between pre-test and post-test scores. It attests to the point that the DA training did help those who received the DA training to enhance their reading comprehension performance.

Table 5. Paired Samples Test

Paired Differences									
95% Confidence Interval of the Difference									
	Mean	Std. Deviator	Std. Error Mean	Lower	Upper	t	df	Sig.(2-tailed)	
Pair1 pre test CO- post test CO	1.07692	1.93484	.53663	-2.24613	.09229	2.007	12	.068	

Table 5 is based on the results of the T-Test for the CG of the study showing the comparison between the pre-test and the post-test mean scores of the CG of the study. They are compared to determine the change level the participants in the group made achieved in their reading comprehension proficiency after static assessment. The T (2.00), degrees of freedom (12), and significance (.068) help reject the null hypothesis that there is no significant difference between the means of the two variables. Here, it is seen that the significance value is statistically significant. That there is difference between pre-test and post-test scores meaning that static assessment did help those who received it to enhance their reading comprehension performance.

Table 6. Independent Samples Test

Levene's Test for Equality of Variances										t-test for Equality of Means									
										95% Confidence Interval of the Difference									
		F	Sig.	t	df	Sig. (2-tailed)	Mean Differences	Std. Error Difference	Lower	Upper									
Pretest Equal variances																			
Assumed		.944	.340	.640	28	.527	.53333	.83305	-1.17309	2.23975									
Equal variances																			
Not assumed				.640	27.567	.527	.53333	.83305	-1.17430	2.24096									

This table consists of two major parts. The first section is related to the equality of the variances of the two samples, but the other one deals with the equality of the means of the two samples. The former is where the Levene's Test for Equality of Error Variances is represented. It tests the null hypothesis that the error variance of the dependent variable is equal across groups. If the Levene's Test is significant (the value under "Sig." is less than .05), the two variances are significantly different. If it is not significant (Sig. is greater than .05), the two variances are not significantly different; that is, the two variances are approximately equal. If the Levene's test is not significant, the second assumption is met. Here, the significance is .000, which is less than .05. Therefore, it is assumed that the variances are significantly different. In this table, the dependent variable is the post-test scores, and the design is Intercept+GROUP\_NA+PRETEST+GROUP\_NA \* PRETEST. Since the Sig. value is below 0.05, the null hypothesis that the variances are equal is rejected and the second section should be studied. As seen in Table 6, T value is .64 and here exists 27.56 degrees of freedom. The Sig. value in the t-value for equality of means equals (.52). Therefore, the null hypothesis the means of the two groups are not significantly different is rejected and it can safely be stated that there is a significant difference between the experimental and control groups in terms of their performance on the post-test on reading comprehension ability. Students who took the DA had significantly higher reading comprehension performance than those who took the Static Assessment.

Table 7. Correlation between pre-test and post-test scores of experimental group

		Pre-test EX	Post-test EX
Pre-testEX	Pearson Correlation	1	.798(**)
	Sig. (2-tailed)		.001
	N	15	13
Post-testEX	Pearson Correlation	.798(**)	1
	Sig. (2-tailed)	.001	
	N	13	13

\*\* Correlation is significant at the 0.01 level (2-tailed).

As it is clear, Table 7 shows that the Sig. value is .001 which is below 0.05, so the H<sub>0</sub> (i.e. there is no significant correlation between the two variables) is rejected and it is hypothesized that there is a significant but not accidental relationship between the pre-test and post-test scores. Correlation is significant at the 0.01 level (2-tailed).

Table 8. Correlation between pre-test and post-test scores of control group

		Pre-testCO	Post-testCO
Pre-testCO	Pearson Correlation	1	.545
	Sig. (2-tailed)		.054
	N	15	13
Post-testCO	Pearson Correlation	.545	1
	Sig. (2-tailed)	.054	
	N	13	13

Here, the table shows that the Sig. value is .054 which is above 0.05, so the H<sub>0</sub> (i.e. there is no significant correlation between the two variables) is accepted and it is assumed that there is a significant, not accidental, relationship between the pre-test and post-test scores. Correlation is significant at the 0.054 level (2-tailed).

#### *4.3. Null Hypothesis 1 Testing*

The main null hypothesis put forth by the researchers was: Dynamic Assessment (DA) does not affect Iranian EFL students' reading comprehension performance. As a result, the first null hypothesis is rejected, and it can be claimed that DA instruction exerts a statistically significant influence on the participants' reading comprehension performance.

#### *4.4. Null Hypothesis 2 Testing*

The second null hypothesis derived from the main null hypothesis was: The experimental group of the study will not show any progress from the pre-test to the post-test of the study. This hypothesis was intended to test the development level the experimental group (N=15) of the research study achieved after DA. The results of the paired t-test helped reject this hypothesis in that the group showed significant development from the pre-test to the post-test of the study. Their reading comprehension gain on the post-test attested to the effectiveness of the treatment they all received for 9 sessions. There was a statistically significant difference between the pre-test and post-test scores.

#### *4.5. Null Hypothesis 3 Testing*

The last null hypothesis of the current study is: The control group of the study will not show any progress from the pre-test to the post-test of the study. In a similar fashion, paired t-test was utilized as an important data analysis method to accept or reject the null hypothesis. The results of the measure rejected this hypothesis, too. The control group of the study also improved their reading comprehension ability at the end of their 9-session treatment of Static Assessment. The major difference between the EG and the CG of the study was the development level they reached after their treatment. In the former, the development was statistically significant while in the latter the reading comprehension gain was not statistically significant, though they showed some progress. To put it another way, the level of significance the EG experienced was statistically higher than that of the CG.

### **5. Discussion**

This longitudinal research study consisting of nine consecutive, 80-minute sessions was to empirically investigate the effects of Dynamic Assessment on the reading comprehension performance of the EG participants in Nosrat Institute, located in Kermanshah. The researchers (mediators) used DA in the EG, while the CG received traditional assessment or Static Assessment. The research instruments included a pre-test (that is, a TOEFL reading comprehension test), and a post-test (that is, a TOEFL reading comprehension test). It should be added that the research was homogenized through an Oxford Placement Test which sorted the population into three groups of reading-low, reading-mid, reading-high students. 30 students were randomly selected from the group with the lowest level of reading comprehension proficiency. The participants in both groups took part in these three consecutive steps of the research procedures: the pre-test (that is, a TOEFL reading comprehension test), DA was used for the EG and SA was used for the CG and the post-test (that is, a TOEFL reading comprehension test). The primary purpose of this study was to determine the effects of utilizing DA Reading Comprehension. To put it another way, the researchers were to determine if DA can be used instead of

SA to improve the learners' Zone of Proximal Development (hereafter ZPD) to enhance the comprehension of content-based lessons.

The results achieved through the data analysis methodologies indicated a trend in the direction of some advantages for the use of DA. Participants in the EG scored higher on the post-test than those in the CG. The study found some statistically significant evidence to support the overall effectiveness of DA use in reading comprehension tasks. A reason for this may have been the long GO training time. The EG had an ample amount of practice and consistent exposure to different types and strategic ways to answer the questions in the reading comprehension texts. This may have resulted in their knowledge gains which were possibly manifested in the post-test scores.

The present study builds on the recommendations of Poehner and Lantolf (2005) and provides empirical support for their claim that DA principles can make classroom formative assessment practices more effective by providing assistance that is continually tuned and returned to learners needs. The inclusion of interaction brought to light the extent of learners' understanding and control over linguistic forms and their relation to meaning, and also helped with the identification of problems underlying poor performance. Evidence was also presented to suggest that interaction provided an opportunity for these problems to be addressed and for learners to develop. The fact that the learners themselves arrived at a greater awareness of their abilities, as evidenced by the verbalization data, is especially important. According to Vygotsky's (1986) argument the goal of instruction should be to render the invisible visible.

In this study, the increase in the post-test performance by the low-reading group was promising. Perhaps this group found DA as an effective learning and/or test performance strategy. This study found statistically significant evidence that DA is effective for the EG. Specifically, DA use has been statistically seen as effective for low-skilled readers. This finding could support previous claims concerning the applying of DA in educational setting in Islamic Republic of Iran. It worth mentioning that the results of the present study confirmed the results obtained in the study of Campione & Brown (1987), Valencia, Campione, Weiner, and Bazzi (1990), Abbott, Reed, Abbott, and Berninger (1997), Kletzien and Bednar (1990), and Kozulin and Garb (2002).

Campione & Brown (1987) were among the pioneers who used the framework of DA to specific educational domains. The principal of their assessment model was to investigate how learners learn from each other and how flexible they were in using what they have learned. Result showed that predictions based on initial performance significantly underestimated what children could achieve with minimal assistance. Therefore, DA measures often provide better estimates of reading comprehension than SA.

Valencia, Campione, Weiner, and Bazzi (1990) applied DA to reading domain (as cited in Haywood & Lidz, 2007, p. 80). They used an experimental dynamic test approach with several control groups and their sample consisted of 196 sixth-grade students. Posttest results showed that increased strategy use and improved reading comprehension for those learners who had been moderately or strongly exposed to scaffolding strategies during intervention. As a matter of fact, this enhancement in strategy use persisted for at least five month. The more intense the scaffolding procedure during the intervention had been, the higher the retest performance at five-month follow-up.

The Abbott, Reed, Abbott, and Berninger's (1997) study of 16 second graders with severe reading and writing disabilities offers yet another approach within the broader DA model. These authors framed their study within the concept of "response to intervention" using growth curves as measures, with their tutorial intervention spanning a full program year. The tutorial targeted orthographic and phonological awareness, accuracy and fluency of word recognition and comprehension monitoring, handwriting fluency, spelling, and composing. The sessions began during the summer with 1-hour twice-weekly meetings over the course of 8 weeks; this continued through the school year with 1-hour once-weekly individual meetings over an average of 36.5 weeks. The lessons were standardized and scripted. This

study showed gains beyond chance for most of the children on most of the measures but also showed their differential responses to the treatment. The results also generated individualized recommendations for their next school year.

The other study was conducted by Kletzien and Bednar (1990) which their clients were fifth-grade students. Initial learning level was established using a reading inventory and learners' reading process and strategies utilization by means of probes and observations. The intervention program consisted of strategy training such as visualization as well as guided and independent practice sessions. Finally, learners were administered a parallel version of reading inventory. The observed increase in reading level was significantly correlated with reading improvements six month after the assessment. Based on the results, they concluded that teachers who instruct students need a "a firm understanding of strategies, their use, and ability to infer strategy use from reader responses" as well as "expertise in utilizing a range of instructional techniques".

The goal of Kozulin and Garb's (2002) study was to explore the feasibility of the development and implementation of the dynamic EFL assessment procedure with at-risk students. The results indicate that the procedure is both feasible and effective in obtaining information on students' learning potential. It was 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. Because of this, they affirmed that the paradigm of dynamic assessment is useful not only in the field of general cognitive performance but also in such curricular domain as EFL learning (Kozulin and Garb's 2002). They continued that "any dynamic assessment that includes an element of intervention depends on the quality of mediation provided by the assessor. In this respect dynamic assessment is closer to a situation of instruction rather than examination (Kozulin and Garb's 2002).

It is noteworthy to mention that the result of this study confirmed the study that conducted by Mardani and Tavakoli (2011). The aim of their study 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. Based on the results, they stressed that "it can be discussed that DA is more than just a formative assessment. 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" Mardani and Tavakoli (2011).

Feuerstein (1998) said, "human beings are open rather than closed systems, meaning that cognitive abilities ... can be developed in a variety of ways, depending on the presence and the quality of appropriate forms of interaction and instruction". In this study, in the students' open systems, the researchers established students' Zone of Proximal Development through the interaction and instruction in a cooperative and equal way with the tool of dialogue.

The results of this study assures us that the EFL reading comprehension integrated with DA can improve students' reading competence and build up their reading confidence in the following chain reflection. Face to face interaction with the teacher shows students' progress to themselves clearly and completely. In the last session, the students were able to answer the questions by themselves so that the researchers could conclude that the DA had a positive effect on reading comprehension.

Based on this study and the other studies, I can say that if we want to use suitable instructional techniques for a better reading comprehension, teachers or rather mediators should apply tasks and feedbacks within dynamic reading-competence tests. Most of the time, DA in the domain of reading

and text comprehension depend on instruction and practice in meta-cognitive knowledge such as strategies which is specific to certain reading tasks and goals. Also, successful DA relies on the prompting of domain-specific processes which are vital to do the tasks and practices.

The results of this study assures us that the EFL reading comprehension integrated with DA can improve students' reading competence and build up their reading confidence in the following chain reflection. Face to face interaction with the teacher shows students' progress to themselves clearly and completely. In the last session, the students were able to answer the questions by themselves so that the researchers could conclude that the DA has a positive effect on reading comprehension.

Based on this study, we can say that if we want to use suitable instructional techniques for a better reading comprehension, teachers or rather mediators should apply tasks and feedbacks within dynamic reading-competence tests. Most of the time, DA in the domain of reading and text comprehension depends on instruction and practice in meta-cognitive knowledge such as strategies which are specific to certain reading tasks and goals. Also, successful DA relies on the prompting of domain-specific processes which are vital to fulfill the tasks and practices.

## 6. Conclusion

The aim of this study was to explore the feasibility and practicality of development and implementation of Dynamic Assessment (DA) to reading comprehension of EFL students in the context of Iran. The present study follows Lantolf and Poehner's (2004) reasoning that DA can optimally promote learner development through application of a clinical methodology, as called for by Vygotsky (1956, 1998) with regard to the Zone of Proximal Development (ZPD), in which mediator and learner collaborate to perform the assessment task. Therefore, this study contributes to the second language DA literature by devising and implementing a DA procedure with 15 institute students in the 16-20 age range. The results of this study displayed that the participants' performance improved dramatically on the post-test after DA- training. As Xiaoxiao and Yan (2010) state, Feuerstein regarded DA as a way of assessing the true potential of children that differs significantly from conventional tests. Dominating the field of language testing, static assessment used to determine whether some pre-determined achievement level had been reached. Traditional static assessment was limited because it did not directly aim to stimulate learners into becoming independent knowledge constructors and problem solvers. Unlike static assessment, DA gives the language teacher a chance to appropriately gauge the students' understanding and ability level and how to improve the students' level development. To put it in another way, by engaging in DA activity, teachers may be able to challenge individuals to reach higher levels of functioning (Poehner, 2005, cited in Naeini and Duvail, 2012).

As a matter of fact DA with its monistic view toward teaching and testing not only assesses the learners' abilities but also provides them with opportunities for learning and development. This in turn has some positive results both for teachers and learners; therefore, the implications can be multifold.

First, it helps students to take the advantage of mediation provided by the assessor and become autonomous in doing similar tasks later on. As we can see the learners were able to answer the reading comprehension questions by themselves in the last session of the study.

Second, applying DA provides the opportunity for learner to being mediated without suffering from stress. In some educational setting such as Iran in which test scores lay the students under stress is a major test score pollution source. Therefore, test scores unaffected by stress factor can be more accurate for educational decisions. So it can be concluded that DA results in presenting a true picture of the abilities which is the most concern of assessment. To put it in another words, DA provides more valid measures of students' ability.

Third, DA of learners' abilities can avoid misinterpretations and misrepresentations of the abilities. Contrary to Static Assessment, DA presents learning potential of the learners because it sheds light on both current zone of the learners and their hidden potential in the zone of proximal development after removing hindering factors.

Fourth, DA intervention in the form of mediation was shown to have a significant role in the diagnostic perspective of students' problem areas. The procedure section provides a lot of examples showing that in many cases that the students were not able to display their second language during the independent performance but they could overcome the problem after receiving implicit or explicit assistance provided by researchers or rather the mediators.

Fifth, this study contributes significantly to provide mediation. As noted by the researchers working within the psychological DA framework, the provision of mediation is one of the most intimidating aspects of DA application in practice (Haywood and Lidz, 2007). This issue has not received much attention in general DA research (Lidz, 1991) as well as in second language DA research (Aljaafresh & Lantolf, 1994; Poehner, 2005). Therefore, the mediational strategies used in this study can offer ways how to apply DA-based mediation.

Sixth, this study introduces a substantial way to assess and teach reading comprehension in empirical research. According to the results, mediational strategies can be useful in diagnosing the source of the problems which second language learners experienced during the reading process.

In the current study, the researchers only worked with 15 male students. Future research can investigate the same research project with both male and female participants and even more participants. This study attempted to use a DA approach in one skill known as reading comprehension. Further research can focus on DA approaches in the other domains such as listening, writing, vocabulary acquisition and so forth.

In this study, students were motivated by the promise of a prize if they "worked hard" and "paid attention." These conditions do not closely resemble whole class instruction in schools. It may be interesting in future studies to investigate whether DA plus a measure of student attention can improve the reading comprehension better than DA alone.

As a matter of fact, different students have different degree of responsiveness to the mediation and the assistance introduced by the mediator, that is the effect of mediation varies from one student to the other student. Future research can work on the extent to which different students benefit from the mediation. Group DA (G-DA) is one of the new directions that recent second language DA research has begun to explore. The future research can focus on this kind of DA and find out what contributions G-DA can have, but in different contexts of more participants to come to a more reliable findings.

## References

- Abbott, S. P., Reed, E., Abbott, R. D., & Berninger, V.W. (1997). Year-long balanced reading/writing tutorial: A design experiment used for dynamic assessment. *Learning Disability Quarterly*, 20(3), 249-263. DOI: 10.2307/1511311
- Aljaafreh, A., & J. P. Lantolf. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 78(4), 465-483. DOI: <http://dx.doi.org/10.1111/j.1540-4781.1994.tb02064.x>
- Bailey, K. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing* 13(3), 257–279. DOI: 10.1177/026553229601300303

- Best, R. M., Rowe, M., Ozuru, Y., & McNamara, D. S. (2005). Deep-level comprehension of science texts: The role of the reader and the text. *Topics in Language Disorders*, 25, 65–83. DOI: 10.1097/00011363-200501000-00007
- Bransford, J. D., Delclos, V. R., Vye, N. J., Burns, M. S., & Hasselbring, T. S. (1987). State of the art and future directions. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 479-496). New York: Guilford Press. DOI: 10.1177/002221948902200303
- Campione, J. C., & Brown, A. L. (1985). *Dynamic assessment: One approach and some initial data* (Technical Report No. 361). Bethesda, MD: National Institute of Child Health and Human Development; National Institute of Education. (ERIC Document Reproduction Service No. ED269735)
- Campione, J. C., & Brown, A. L. (1987). Linking dynamic assessment with school achievement. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 82–140). New York: Guilford Press.
- Cioffi, G., & Carney, J. J. (1983). The dynamic assessment of reading disabilities. *Reading Teacher*, 36, 764–768.
- Ellis, R., (2000). Task-based research and language pedagogy. *Language Teaching Research*, 4(3), 193-220. DOI: 10.1177/13621688000400302
- Feuerstein, R. (1998). *The theory of mediated learning experience: About human as a modifiable Being*. Jerusalem: Ministry of Defense Publications (in Hebrew).
- Feuerstein, R., Rand, Y., & Hoffman, M. B. (1979). *Dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques*. Baltimore, MD: University Park Press.
- Hamers, J., Pennings, A., & Guthke, J. (1994). Training-based assessment of school achievement. *Learning and Instruction*, 4, 347–360.
- Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice: Clinical and educational applications*. Cambridge, UK: Cambridge University Press.
- Jitendra, A. K., & Kameenui, E. J. (1993). An exploratory study of dynamic assessment using two instructional strategies on experts and novices' performance in solving part-whole mathematical word problems. *Diagnostique*, 18(4), 305-324. DOI: 10.1177/073724779301800404
- Karpov, Y., & Haywood, H. C. (1998). Two ways to elaborate Vygotsky's concept of mediation: Implications for education. *American Psychologist*, 53(1), 27-36. DOI: 10.1037/0003-066X.53.1.27
- Kletzien, S. B., & Bednar, M. R. (1990). Dynamic assessment for at-risk readers. *Journal of Reading*.
- Kozulin, A. and E. Garb. (2002). Dynamic assessment of EFL text comprehension of at-risk students. *School Psychology International* 23, 112–127. DOI: 10.1177/0143034302023001733
- Lantolf, J.P. (Ed.). (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.
- Lantolf, J.P., and Appel, G., (1994). Theoretical framework: an introduction to Vygotskian approaches to second language research. In: Lantolf J. P., ed. *Vygotskian approaches to second language research*. London: Ablex Publishing, 1-32
- Lantolf, J.P. and M.E. Poehner. (2004). Dynamic assessment: bringing the past into the future. *Journal of Applied Linguistics* 1: 49–74
- Lidz, C.S. (1991). *Practitioner's guide to dynamic assessment*. New York: Guilford.

- Lunt, I. (1993). The practice of assessment. In H. Daniels (Ed.), *Charting the agenda: Educational activity after Vygotsky*, 145-170. London and NY: Routledge.
- Mardani, M. and Tavakoli, M. (2011). Beyond reading comprehension: The effect of adding a dynamic assessment component on EFL reading comprehension. *Journal of Language Teaching and Research*, Vol. 2, No. 3, pp. 688-696. Retrieved from <http://www.ojs.academypublisher.com/index.php/jltr/article/viewFile/0203688696/3066>
- María del Carmen Malbrán; Claudia M. Villar. (2002). The dynamic evaluation of the learning potential: a procedure. *Orientación y Sociedad*, 3, 1-12. Retrieved from [http://www.scielo.org.ar/pdf/orisoc/v3/en\\_v3a09.pdf](http://www.scielo.org.ar/pdf/orisoc/v3/en_v3a09.pdf)
- McNamara, T. (2004). Language testing. In *The Handbook of Applied Linguistics* (pp. 763–783). A. Davies and C. Elder (Eds.). Malden, MA: Blackwell.
- Oxford Placement Test. [http://www.lang.ox.ac.uk/courses/tst\\_placement\\_english.html](http://www.lang.ox.ac.uk/courses/tst_placement_english.html)
- Palincsar, A., & Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117-175.
- Poehner, M. (2005). Dynamic assessment of oral proficiency among advanced L2 learners of French. (Unpublished doctoral dissertation). The Pennsylvania State University, the U.S. Retrieved from <https://etda.libraries.psu.edu/paper/6627/1890>
- Poehner, M.E. and J.P. Lantolf. (2005). Dynamic assessment in the language classroom. *Language Teaching Research* 9 (3). DOI: 10.1191/1362168805lr166oa
- Seng, Alice S.H. , Lucy Pou & Tan Oon Seng. (Eds) (2003). *Mediated learning experience with children*. Applications Across Contexts. Singapore: McGrawHill
- Tissink, J., Hamers, J. H. M., & Van Luit, J. E. H. (1993). Learning potential tests with domain-general and domain-specific tasks. In J. H. M. Hamers, K. Sijtsma, & A. J. J. M. Ruijsenaars (Eds.), *Learning potential assessment: Theoretical, methodological and practical issues* (pp. 243–266). Amsterdam: Swets & Zeitlinger.
- Valencia, S.W., Campione, J. C., Weiner, S., & Bazzi, S. (1990, December). *Dynamic assessment of reading achievement*. Paper presented at the National Reading Conference, Miami, FL.
- Vygotsky, L. S. (1995). Apprentissage et développement à l'âge préscolaire. *Société française*, 2/52, 35-46.
- Vygotsky, L.S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1979). Consciousness as a problem in the psychology of behavior. *Soviet Psychology*, 17, 3-35.
- Vygotsky, L. S. (1986). *Thought and language*. (A. Kozulin, trans.). Cambridge, MA: MIT Press.
- Vygotsky, L.S. (1997). *The crisis in psychology*. In *The Collected Works of L. S. Vygotsky*. Vol. 3. *Problems of the Theory and History of Psychology Including the Chapter on the Crisis in Psychology*. R.W. Rieber (Ed.). New York: Plenum.
- Vygotsky, L. S. (1998a). Infancy (M. Hall, Trans.). In R. W. Rieber (Ed.), *The collected works of L. S. Vygotsky*: Vol. 5. *Child psychology* (pp. 207-241). New York: Plenum Press. (Original work written 1933-1934)
- Vygotsky, L. S. (1998b). The problem of age (M. Hall, Trans.). In R. W. Rieber (Ed.) Weber, R. P. (1990). *Basic content analysis* (2nd ed.). Newbury Park, CA. Sage.
- Wertsch, J., 1985. *Vygotsky and the social formation of mind*. UK: Harvard University Press.

Williams, M. and Burden, R., (1997). *Psychology for language teachers, a social constructivist approach*. Cambridge: Cambridge University Press.

Xiaoxiao, L., & Yan, L. (2010). A case study of dynamic assessment in EFL process writing. *Chinese Journal of Applied Linguistics*, 33(1), 24-40.

## Dinamik değerlendirmeye karşı statik değerlendirme: İngilizceyi yabancı dil olarak öğrenen İranlılardaki okuduğunu anlama yetisi üzerine bir çalışma

### Öz

Bu çalışma, yabancı dil bağlamlarında okuma yetisinin geleneksel veya statik değerlendirmesini verir ve öğrencilerin okuma yetisinin gelişiminde dinamik değerlendirmeyi (DA) kullanır. Araştırma evrenini (N=250) bağdaşık hale getirmek için bir Oxford seviye tespit sınavı (OPT) yapılmıştır. Sınav sonuçlarına dayanarak evren; okuma düzeyi düşük, orta ve yüksek öğrenciler şeklinde üç gruba ayrılmıştır. Bu çalışmanın katılımcıları; kontrol ve deney gruplarına gelişigüzel ayrılan, okuduğunu anlama yetkinliği en düşük düzeyde olan 30 katılımcısı olan bir örneklemdir. Bu çalışma 2012 yılında Kermanshah şehrinde Nosrat Enstitüsünde gerçekleştirilmiştir. Mevcut eğitim alanlarını ve okuma yetkinlik düzeylerini belirlemek için iki gruba da okuduğunu anlama ön testi yapılmıştır. Ardından dinamik değerlendirme (DA) çalışması; deney grubuna (EG) okuduğunu anlama üzerine 80 dakikalık birbirini takip eden 9 oturumda uygulanmıştır, deney grubunun aksine kontrol grubuna (CG) da statik değerlendirme uygulanmıştır. Sonunda eğitimden sonraki okuduğunu anlama başarı düzeyini ölçmek için araştırma grubuna başka bir TOEFL okuduğunu anlama ön testi yapılmıştır. İstatistiksel veri analizi, dinamik değerlendirmenin düşük yetiye sahip okuyucular için statik değerlendirmeden daha önemli ve etkili olduğunu göstermiştir.

*Anahtar Sözcükler:* Değerlendirme; dinamik değerlendirme; statik değerlendirme; yakınsal gelişim alanı; okuduğunu anlama yetisi

### AUTHORS' BIODATA

**Behzad Nazari** has an M.A. in TEFL from Islamic Azad University, South Tehran Branch. His areas of interest are Sociolinguistics, Discourse, and Dynamic Assessment. He has been teaching English as a foreign language in Iran for many years and he is an English teacher at the ILI.

behzad.nazari1362@yahoo.com      b.nazari62@gmail.com

Address: Iran Language Institute, Rasht St., Valiaasr St., Tehran, Iran.

Telephone Number: 0098 935 681 6340

**Saeed Mansouri** has an M.A. in TEFL from Islamic Azad University, Tonekabon Branch. He is interested in Assessment, Teaching theories, and the related fields. He has been teaching English as a foreign language in Iran. Saeed.mansouri1360@gmail.com