



Content list available at www.urmia.ac.ir/ijltr

*Iranian Journal
of
Language Teaching Research*



Urmia University

Exploring transcendence in EFL learners' reading comprehension through computerized dynamic assessment

Saman Ebadi^{a,*}, Abdulbaset Saeedian^a

^a *Razi University, Iran*

ABSTRACT

Derived from Vygotsky's works, dynamic assessment (DA) enables learners to move beyond their current level of functioning through offering needs-sensitized mediation. This study aimed at exploring the learners' development in novel and increasingly more challenging situations called transcendence (TR) in an L2 context focusing on reading comprehension through computerized dynamic assessment (CDA). An overall number of 32 BA TEFL advanced students were selected from among undergraduates of a university in Iran to participate in this study. To fulfil the purpose of the study, the researchers developed Computerized Dynamic Reading Assessment to examine Transcendence (CDRAT) software which included reading comprehension CDA tests. To monitor the maintenance of learning in different time intervals, two software programs named CDRAT1 and CDRAT2 were utilized in specific time intervals after the posttest. The results indicated that not only did increased task complexity not lead to regression of students' development, but also it had an effect on enhancing their development. Thus, transcendence had an effect on the generalizability of the contentions to a great extent. In conclusion, it was disclosed that there is no endpoint to progress and merely gaining a low or high score should not be interpreted as the inability of a special learner to surpass. This would more reinforce the need for applying TR in other future DA studies.

Keywords: dynamic assessment; computerized dynamic assessment; transcendence; reading comprehension; mediation

© Urmia University Press

ARTICLE HISTORY

Received: 1 Jul. 2014

Revised version received: 19 Dec. 2015


Accepted: 21 Dec. 2015

Available online: 1 Jan. 2016

* Corresponding author: English Language Department, Faculty of Humanities, Razi University, Kermanshah, Iran

Email address: samanebadi@gmail.com

© Urmia University Press

 10.30466/ijltr.2016.20376

Introduction

There are too many studies whose results indicate that instruction and assessment are generally viewed as separate and in some cases opposite activities (Linn, 2000; Lynch, 2001; McNamara, 2001; Moss, 1996; Shohamy, 1998, 2001). Seeking to find an answer for the relationship between assessment and instruction, the researchers attempted to apply an alternative method on assessment and its relevance to teaching and learning based on the sociocultural theory (SCT) of mind developed by L.S. Vygotsky. In Vygotskian SCT, man's cognition is studied within a social context and mediated by symbolic and physical tools (Azabdaftari, 2013 a). Dynamic assessment (DA) which is framed within the SCT contends that assessment and instruction are not separated but instead are fully integrated by challenging the conventional views concerning these two terms (Poehner, 2008). The integration, based on Lidz and Gindis (2003), occurs since during the assessment procedure mediation or intervention is offered to enhance learners' functioning. Like any other methods, some drawbacks such as being time consuming, focusing on low number of participants, etc. lie within DA, as well. To overcome these shortcomings, a type of DA called computerized dynamic assessment (C-DA) was called upon.

There exists a robust literature focusing on DA (Kozulin & Garb, 2002; Lidz & Gindis, 2003; Poehner, 2005, 2008), on DA in reading comprehension (Kozulin & Garb, 2001; Shabani, 2012.), on C-DA in reading (Pishghadam & Barabadi, 2012; Teo, 2014, 2012) but few studies (Birjandi & Ebadi, 2012; Ebadi, 2014; Poehner, 2008; Poehner & Lantolf, 2013), to date, have taken transcendence (TR) into account. TR is regarded as an important concept in DA (Poehner, 2008) since it emphasizes the need for variable contexts in preference to homogeneous ones and in TR, the development of learners from one DA interaction to another novel one is of significance. Thus, the present research has applied the use of C-DA on reading comprehension while focusing on TR.

The core focus of this study was the concept of transcendence which emphasizes variable contexts rather than homogeneous ones, presents learners with increasingly complex problems and gives careful attention to their performance. The concept of *transcendence* was first introduced by Feuerstein, Rand, and Rynders (1988) in the famous list of Mediated Learning Experience (MLE) attributes to underscore that development extends well beyond any given task. Thus, this study attempted to find out if the students were able to sustain their improved performance in increasingly more challenging and novel assessment contexts. TR was included in this study as the researchers attempted to ensure about tracking learners' development (sustenance, progression, or regression) from one DA interaction to another and also they aimed to avoid generalizing individuals' performance on posttest to other future contexts.

Literature Review

Computerized Dynamic Assessment (C-DA)

Similar to DA, the central concept of the computerized dynamic assessment (C-DA) is grounded in Vygotsky's theoretical framework (1978). There are two approaches to CDA, namely *interactionist* and *interventionist*. The interactionist approach to CDA (Birjandi & Ebadi, 2012) is against conducting quantitative research on the area of dynamic assessment and embraces a qualitative approach. The mediation which is provided for different individuals is attuned to their zone of proximal development (ZPD) in which the movement towards the future is uncertain, open, and mediated (Azabdaftari, 2013 b). In the interventionist approach to DA, hints and prompts are pre-specified. With regard to its advantages, Poehner and Lantolf (2010, p.318) indicated

“Interventionist approaches to DA have the advantage of efficiency because they can be simultaneously administered to large cohorts of individuals, especially in computerized format.”

So far a few known interventionist DA and C-DA studies have been conducted in the field of education in the L2 context; therefore, what follows is the applications of C-DA principles to L2 learning contexts.

To find out the success of C-DA in improving students' performance of reading comprehension and also to determine whether the difference between means of dynamic and non-dynamic scores has been statistically significant, Pishghadam and Barabadi (2012) utilized the t-test and through citing Farhadi, Jafarpur and Birjandi (1994), they concluded that their test had moderate reliability and validity. With regard to construct validity, they cited Poehner (2007) and Lidz and Macrine (2002) who contended that a DA test has construct validity if it results in significant improvement of test takers' performance. Pishghadam and Barabadi concluded that based on the results of their study C-DA had helped test takers improve their scores which, in turn, ensures construct validity of CDA test.

The findings of Pishghadam and Barabadi's (2012) study highlighted the effectiveness of C-DA in making students' reading comprehension ability better and in getting information about the participants' learning potentiality which is different in a better way from their initial performance level. One of the shortcomings of their study is that it had not held any transfer sessions. In more technical words, Feuerstein's attribute of transcendence which is one of the most important attributes in transforming a given interaction into a mediated learning experience (MLE) and is obligatory for DA to fully integrate assessment and instruction (Poehner, 2005) has not been queried in their study. Poehner (2005) contends that it is necessary to hold TRs to claim that a change in the participants' performance has happened through DA procedures.

Tzuriel and Shamir (2002) carried out a study in which mediation provided by computer (C-DA) and mediation provided by human mediator (DA) were queried and compared. The value of mentioning this study here is twofold: first, it was carried out in a L2 context and second and like our study, it examined the effect of mediation provided by the computer. Having selected a sample of 60 kindergarten children, Tzuriel and Shamir assigned them into two groups with equal number of participants: Computer Assisted (CA, n= 30) and Examiner Only (EO, n= 30). The results of the study indicated that while both groups improved their gain scores from pretest to posttest, the outperformance of the CA group was significantly obvious. At the very end of the study, they contended that without the presence of a human mediator it was impossible for the computerized mediation to become successful.

The studies done by Tzuriel and Shamir (2002) and Pishghadam and Barabadi (2012), have one problem in common. Transcendence has not been examined by them. In this regard, Poehner (2005, pp.111-112) stated the “view of learning as transcending the assessment means that the mediator's goal is not simply to get the learner through the task at hand, but rather to use their interaction as an opportunity to understand and promote the learner's abilities.” This means that in transcendence, it is very important to understand the relationship between performance and context which is conceptualized by DA to make sure if learners' performance in variable contexts can be as good as their performance in homogeneous ones or not. In the present study, every attempt has been made to overcome this shortcoming by holding two Transfer Assessments (TR) after the posttest; TR1 and TR2. These two TRs have been taken after the posttest in specific time intervals to monitor the maintenance of learning in different time intervals. The impact of DA on learners' development, being underscored by Poehner (2008) as well, cannot be appreciated without taking the theoretical constructs including transcendence. The following research question guided the study:

To what extent, if the learners progress, will they be able to sustain their improved performance in transcendence tasks?

Methodology

Design

The design of this study included the following stages: pretest, enrichment program, posttest, and transfer assessments. Through these stages three scores were obtained: actual or NDA score (indicating participants' unmediated performance); mediated or DA score (showing the degree of the participants' receptiveness to help provided for them on each test item); and a learning potential score (representing the difference between their actual and mediated performance).

This study which followed DA principles (Ableeva, 2010) and interventionist approach to DA (Poehner & Lantolf, 2005) employed both qualitative and quantitative methods of data analysis. In other words, based on Poehner and Lantolf's (2005) claim that interventionist DA is rooted in quantitative interpretation of ZPD, the present study is quantitative because it tracks the paradigms of interventionist approach to DA.

Participants

An overall number of 32 BA TEFL advanced students were selected from among undergraduates of a University in Iran to participate in this study. The mean of their age was 27 meaning that they could be regarded as adults. This study was conducted in an EFL context as English was the foreign language of these adult participants. Concomitantly with Poehner (2005), the proficiency of the learners was determined by the number of semesters spent in university. Since they had already passed eight semesters, they were advanced students. Besides, the results from the DIALANG, a free online assessment system to determine individuals' proficiency level, were also used to confirm the homogeneity of the participants. But among the 32 participants, the results showed that 24 were at the B2 English reading comprehension level, 7 were at the B1 proficiency level, and only one participant was at the C1 level representing the learner's being at-risk.

Instruments

The data collection means consisted of the following instruments: the DIALANG software, the Computerized Dynamic Reading Test (CDRT) developed by Pishghadam and Barabadi (2012), and two Computerized Dynamic Reading Assessment to examine Transcendence (CDRAT) developed by the present researchers. Needless to say, this is the first time that such CDRATs are developed to investigate transcendence. That is, in most of the other studies done in this area, transcendence has either not been taken into consideration (e.g. Pishghadam & Barabadi, 2012; Tzurriel & Shamir, 2002) or it's not been computerized (e.g. Ableeva, 2010).

DIALANG

DIALANG is a free online assessment system available at <http://dialangweb.lancaster.ac.uk> which provides learners with sufficient information about their linguistic proficiency. In addition, it also shows the full feedback (i.e. results and advice,) and a review of the responses to the items showing how well they have taken the test upon completion of the test. It determines individuals' proficiency levels based on the levels from the Common European Framework of Reference for Languages

(CEFR) starting from A1 (Breakthrough) and A2 (wastage) as Basic User through B1 (Threshold) and B2 (Vantage) as Independent User to C1 (Effective Operational Proficiency) and C2 (Mastery) as Proficient User.

Readability Formulae

To strengthen the claim that the passages used for the CDRATs were more challenging, the researchers used the readability formulas available at the website: <http://www.readabilityformulas.com/> which is a free website to help readers: 1) score their texts (documents, books, policies, technical materials, etc.) and 2) find the reading level and grade level that different readers need to read and comprehend their texts. The purpose of Text Readability Consensus Calculator which uses 7 popular readability formulas (a combination of all) is to calculate the average grade level, reading age, and text difficulty of the sample text which is pasted in that website. In other words, it helps writers and non-writers make sure if their readers can understand what they write through evaluating their writings in this way.

With regard to the usefulness of these formulae, it should be stated that their first use, to the best knowledge of the researchers, dates back to Flesch (1948). In the past, these formulae had been used over time by many researchers such as Farr, Jenkins, and Paterson (1951), and Pichert and Elam (1985), to name just a few. This does not mean that these formulae are not used any more but instead they are still regarded as guidance for the development and evaluation of literacy-based written texts (Wang, Miller, Schmitt, & Wen, 2013). The most recent source regarding the importance of the readability formulae belongs to Hernández-Murillo and Shell (2014) who specifically focused on Federal Open Market Committee (FOMC) which is one of the formulae used in 1994.

Computerized Dynamic Reading Assessment to Examine Transcendence (CDRAT)

Computerized Dynamic Reading Assessment to examine Transcendence (CDRAT) aims at assessing the participants' English reading proficiency level in novel and more challenging contexts; it aims at measuring one of the important issues in this study, i.e., transcendence. To do so, two software programs named CDRAT1 and CDRAT2 were developed by the researchers. Each one consists of two reading passages and the total number of question items is the same as that of the software which was used in the posttest of the study, i.e. twenty questions for each software.

Unlike Pishghadam and Barabadi's (2012) CDRT in which the hints were only based on some books, the hints prepared for the transcendence stage of the present study were based on the analysis of the results of the pretest, posttest, EP sessions in DA and more importantly on two-stage piloting. A total of 4 learners participated in the pilot phase. Upon their agreement, the first piloting was done as the CDRAT tests were being designed. This was administered traditionally (i.e., without provision of any mediation) in paper-pencil format to achieve a two-folded aim: first, to reassure about the intended skills and second, to gain an overall measurement of the test difficulty along with the specific individual items. The findings were then applied to revise the tests.

The second piloting, however, was administered on the basis of one-on-one interactions dynamically (i.e. with provision of hints on the part of the mediator) albeit in paper-pencil format anew. As clear, the second piloting utilized an interactionist approach to DA, which, as Ableeva (2010, p. 260) holds, "privileges mediation determined on the basis of mediated dialoguing between the teacher and the learner during dynamically conducted assessments." This resulted in strengthening and standardizing the prompts or hints that were used in the CDRATs as they were made on the basis of dialogic engagement through interactionist DA to meet the purpose of

standardization of mediation in interventionist DA. In other words, as in Poehner, Zhang and Lu's (2014) study, the prefabricated mediating prompts were indicative of interventionist DA but the prompts were actually extracted from the interactionist DA administration of the second pilot test. In addition, during process of administering the TRs, teachers were also available and free to provide additional assistance in case required; this resembles interactionist DA as well.

Due to inclusion of TR as a feature which all DA studies have in common (Lantolf & Poehner, 2013), the passages used in the CDRATs had higher difficulty level compared to the ones used in Pishghadam, Barabadi, and Kamrood (2011) and Pishghadam and Barabadi's (2012) CDRT. Even comparatively, the passages in CDRAT2 were more challenging than the ones in CDRT1 based on Readability Formulae although they all assessed the same constructs or skills as the rest of the test.

CDRATs utilized two passages which were more difficult than the ones used in the CDRT. Each passage comprised of 10 items (thus a total of 20 items) which were the same in number as the CDRT and reassessed the same constructs or skills measured too as in Lantolf and Poehner's (2013) study. The researchers developed the program in a way that it easily runs, provided that Net Framework 3.5 is previously installed on the systems. It stores individuals' responses separately, and while recording their correct and incorrect responses it generates a scoring file containing both "mediated and non-mediated (i.e. independent) scores" as in Poehner and Lantolf (2013, p.328). The unmediated score is what the students would have gotten in a traditional test where only the first responses counted, but the mediated score is the weighted score that includes points earned for attempts beyond the first.

The passages which reflected the same nine skills of the CDRT but in a more challenging context were adapted from TOEFL tests but to align with the proposed items in the CDRT, the items were designed in a way that minimized, to a large extent, the possibility of earning the correct response by guessing upon provision of mediation. In this regard, because of the nature of the DA approach which allows individuals multiple chances to answer each item by providing mediation, the number of distractors was increased to at least 6 per item. A drawback of the CDRT which was removed in the CDRATs regarded those individuals who respond correctly to an item at the first attempt; they were not provided with any explanation to minimize the possibility of guessing the correct response.

Arranging the prompts from the most implicit to the most explicit (Aljaafreh & Lantolf, 1994) was another step at preparation of the software programs. In general, individuals were provided with five gradually progressing levels of mediation endeavoring to assist them find the correct response. At first and in line with Pishghadam and Barabadi (2012), they were given 20 minutes to read the whole passage and at the same time the first item was shown on the screen too but as soon as they started answering the first item, they had to come to an answer within a period of 4 minutes. Each item was worth 5 points, hence for each attempt beyond the first individuals were penalized by 1 minus point and in case they could not respond within the specified time, they missed the item and their scores would be recorded as zero (0).

However, if an individual failed to respond correctly to an item at the first attempt, s/he was then given another opportunity to try that item again with the most implicit mediating hint. In case his/her second attempt was incorrect as well, another hint which is more explicit (compared to the first one) was offered. This process kept on and individuals were penalized for each incorrect attempt until they could respond correctly or they used all five attempts and finally the solution and subsequent explanation were shown. A drawback of the CDRATs is when individuals respond correctly at the second attempt; they are not offered any explanation indicating why it is the correct

response before moving to another item. To illustrate how the hints are offered automatically, an example of the CD RAT1 has been addressed below.

Reading Passage	Question and Item Choices		
<p>Paragraph 1</p> <p>Icebergs are massive blocks of ice, irregular in shape; they float with only about 12 percent of their mass above the sea surface. They are formed by glaciers—large rivers of ice that begin inland in the snows of Greenland, Antarctica, and Alaska—and move slowly toward the sea. The forward movement, the melting at the base of the glacier where it meets the ocean, and waves and tidal action cause blocks of ice to break off and float out to sea.</p>	<p>Question 1. According to paragraph 1, all of the following are true of icebergs EXCEPT Write the appropriate letter [A-F] in the following box and then press OK.</p> <table border="1" data-bbox="674 445 877 477"> <tr> <td style="width: 50px; height: 15px;"></td> <td style="width: 50px; height: 15px; text-align: center;">OK</td> </tr> </table> <p>(A) They do not have a regular shape. (B) They are formed where glaciers meet the ocean. (C) Most of their mass is above the sea surface. (D) Waves and tides cause them to break off glaciers. (E) They don't move towards the sea in a high speed. (F) Large rivers of ice are actually not icebergs.</p>		OK
	OK		

Figure 1. Sample Reading Paragraph from the CD RAT1 Passage and its Accompanying Test Item

As figure 1 depicts, the sample paragraph has been derived from a longer passage and here the item asks about the incorrect information on icebergs. The true response is: (C) Most of their mass is above the sea surface. This item targets on the eighth skill, i.e. Factual Information or Stated Detail Questions. If an individual responds correctly at the first attempt, he/she will be shown the following explanation on why it is the correct answer: *“The answer lies in the first line where some information about ‘the sea surface’ has been written. As clear, the phrase ‘only about 12 percent’ unfolds that about 88 percent of icebergs mass is NOT above the sea surface. So, choice c does not provide true information about icebergs.”* A nuanced description of other mediations (in case of choosing the incorrect choice) is as follows.

Mediation 1. In case individuals’ first attempt is incorrect, they are assisted by the first scripted and most implicit hint: *“Sorry, your answer is wrong. Please try again.”* Unlike Teo (2012) who offered a burdensome, long and the most EXPLICIT prompt in our viewpoint, the mediation which was offered in our study for ‘Mediation 1’ was the pilot groups’ recommendations (derived from one-on-one DA interactions) about maintaining the passage visible during the assessment process which was found to be enough opportunity for those who failed at the first attempt as in Poehner et al.’s (2014) explanation.

Mediation 2. If individuals fail to respond correctly again, they are offered more explicit mediation compared to the previous one: *“As it is clear, the question is about “icebergs”, so you should scan the paragraph to find this word and take note of whatever is true about it.”* This hint limits the search zone only to the specified word and leads individuals to *‘take note of whatever is true about’* icebergs.

Mediation 3. Individuals’ failure for the third time buttresses that they require another even more explicit prompt (relative to the previous ones) so that they are offered: *“The information which you have taken note of about ‘icebergs’ should be checked in the answer choices and eliminate those choices which are true about ‘icebergs’ based on the passage. For instance, you can check whether ‘icebergs’ have a regular or irregular shape, or even see if place is important in forming them, etc.”* some examples have been given to direct them concerning how to eliminate distractors based on their notes about ‘icebergs’.

Mediation 4. In the event that individuals cannot come to a conclusion about the correct response anew, mediation is still available and it is even more increasingly explicit as it concentrates on only

ONE sentence, ONE phrase, or even a single word. Anyway, the fourth hint is: *“Be careful about the use of the words “only” and “about 12 percent” in the paragraph, because they may sometimes cause misunderstanding. Now check if you can find wrong information about icebergs.”* As it is clear, individuals are pinpointed to the exact sentence which contains the answer but they should also have kept in mind from the EP sessions in DA that in Factual Information or Stated Detail Questions, sentences are reworded and that though most of the time the correct response is expressed as the same idea which is written in the passage, the words are not exactly the same and instead they are closest in meaning to that presented idea.

Mediation 5. The mediation offered here is the most explicit one and an individual who arrives at this stage obviously reveals that he/she is still not mature enough for the targeted construct or skill which has been taught or in Teo’s (2012, p.14) words he/she “is still far from fully mastering the reading strategy required to understand the concept tested.” Though it is the last offered prompt whereby the correct response is revealed, individuals are still provided with more explanation representing a justification for why that choice is true. The last hint in all CDRATs items is the same: “The true answer is ...” (Here it is choice C) but they are provided with item-specific explanations of the correct response. The explanation for the item under investigation is: *“The answer lies in the first line where some information about ‘the sea surface’ has been written. As clear, the phrase ‘only about 12 percent’ unfolds that about 88 percent of icebergs mass is NOT above the sea surface. So, choice c does not provide true information about icebergs.”*

In conclusion, even though for all of the test items the first and fifth prompts were the same and in each case an explanation on the correct response was given, the other three hints were not fixed, i.e. they were strategy-based and dependent on the targeted skill in different items. As a matter of fact, the three hints were derived from one-on-one DA-based pilot interactions with the 4 learners.

CDRAT1 (TR1)

The data collection procedure at this stage was initiated one week after the posttest and it was dedicated a one-week period to collect all data from TR1. Anyway, to ensure the passages had a higher complexity level in comparison to the ones used in the CDRT, the research-based readability formulae, described earlier, were used for each passage. First, the CDRT passages were calculated and then the results of the passages in CDRAT1 were reported. The CDRT passages were of almost the same difficulty level as follows.

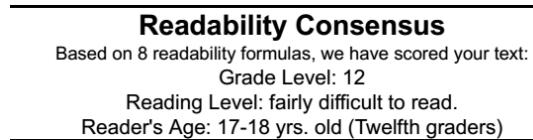


Figure 2. The Overall Consensus of the Passages in the CDRT (Pishghadam & Barabadi, 2012)

Figure 2 shows the overall calculation of the CDRT passages used by Pishghadam and Barabadi (2012) taking the 8 readability formulae into consideration. The results show that the passages are just ‘fairly difficult to read’ and that they are suitable for individuals of 17 to 18 years of age.

Readability Consensus
 Based on 8 readability formulas, we have scored your text:
 Grade Level: 13
 Reading Level: difficult to read.
 Reader's Age: 18-19 yrs. old (college level entry)

Figure 3. The Overall Consensus of the First Passage Entitled 'Green Icebergs' in the CDRAT1

As illustrated in figure 3, the first passage used in the CDRAT1 has a higher difficulty level in comparison to the CDRT ones; an indication for its suitability to be used to gain the purpose of TR in the present study. It is calculated to be difficult to read and appropriate for college level entry. Comparatively, the second passage of the CDRAT1 has higher difficulty level than the first one, too. This means that the way they are ordered has been taken into important consideration as well.

Readability Consensus
 Based on 8 readability formulas, we have scored your text:
 Grade Level: 14
 Reading Level: difficult to read.
 Reader's Age: 21-22 yrs. old (college level)

Figure 4. The Overall Consensus of the Second Passage Entitled 'Desert Formation' in the CDRAT1

As explained, figure 4 shows that the difficulty level of the second passage of the CDRAT1 is more than the first one owing to its appropriateness for readers of older age and level (21-22 years old and grade level, 14 respectively).

CDRAT2 (TR2)

CDRAT2 to which a period of one week was assigned to complete the data collection procedure began 3 weeks after the posttest. In other words, the data collection procedure in TR2 started two weeks after TR1. Needless to say, the passages used in TR2 were the most complicated ones compared to the others used in the study.

Readability Consensus
 Based on 8 readability formulas, we have scored your text:
 Grade Level: 15
 Reading Level: difficult to read.
 Reader's Age: College graduate

Figure 5. The Overall Consensus of the First Passage Entitled 'Children and Advertising' in the CDRAT2

The figure shows that the text which was used in the first passage of the CDRAT2 was more difficult than all used texts in the study so far because it is 'difficult to read' and appropriate for 'college graduates' (the participants of the study).

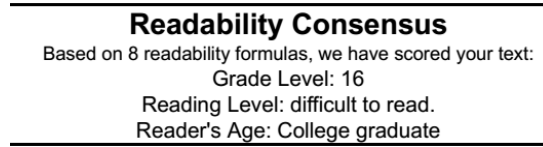


Figure 6. The Overall Consensus of the Second Passage Entitled 'Early Cinema' in the CDRAT2

Though it seems that the second passage is exactly as difficult as the first CDRAT2 passage, it is actually more difficult owing to its grade level (16). This is the most difficult passage of all, that's why it was placed here as the second passage of the CDRAT2. Therefore, the claim on having more difficult and challenging passages in the TR is achieved through research-based readability formulae.

Pretest

The researchers relied on the feedback provided from the DIALANG results and selected the two passages while taking account of both the posttest passages and the areas in which their items concerned about as well. Through taking the paper and pencil reading comprehension test, the participants' actual or non-mediated score was obtained. Needless to say, the students were not allowed to consult each other while taking the test, i.e. they had to take the test individually without making use of any mediation on the mediators' (researchers') side and/or on the other examinees'.

The Enrichment Program: Procedure

Based on the participants' pretest results, a period of two weeks was determined for the Enrichment Program (EP) in DA in a way that each week two tutoring sessions were held for one and a half hours. Contrary to the other three stages of this study, all four EP sessions were conducted in a group format and because of not having a spacious class for 32 people, the participants were divided into two groups (each group containing 16 students) based on some factors such as the time when they were free, etc. which they themselves determined.

The pretest and DIALANG results indicated that the main areas of the participants' problems were lexical, grammatical, and psychological and since the non-computerized EP sessions were held in a group format, the points where all of the learners exhibited problems were given special attention to get them overcome.

Though a lot of areas could be regarded as problematic, just the most important ones were highlighted here. The areas where the participants showed more problems in based on the results of the pretest were the major considerations of the EP sessions which were held in DA. The problems were:

1. The learners' inability in connecting the ideas in the passages.
2. Their confusion about determining the meaning of vocabularies or words.
3. Their difficulty in distinguishing minor or least important details from the significant information.
4. The impact of external factors on their performance. Some external problems such as lack of concentration or distraction during reading, the novelty of this type of assessment for them, etc. could also be considered as the problems they encountered throughout the process.

Determining participants' problems during pretest and providing mediation for them does not, by any means, guarantee the effectiveness of the program. In other words, according to Aljaafreh and Lantolf (1994), the usefulness of corrective feedback is highly dependent upon the nature of the transaction and mediation provided by the expert in this procedure. Therefore, to make the description of the mediator-learner interventions during the EP sessions systematic and to organize the kinds of mediation which are to be used, (Poehner, 2005) 'mediation typology' was used. This typology moves from implicit to explicit forms of intervention and provides the opportunity for the researcher to analyze learners' intervention while taking the quality and frequency of mediation into account. Furthermore, to take the responsiveness of learners to mediation into consideration, the *reciprocity* typology which refers to the extent to which the learners show the ability to take responsibility for their own performance Poehner (2005) was used.

To monitor the progress made from other-regulation to self-regulation, the level of internalization from other-regulation to self-regulation functioning developed by Aljaferh and Lantolf (1994) was used. These five levels of mediation strategies which are transitional were actually developed to evaluate mediation within the ZPD and to keep track of learners' microgenetic development within the C-DA program and transfer assessment tasks. These levels of strategy intervention have been used in some studies to assess learners' language progress (Bijandi & Ebadi 2010a; Oskoz, 2005).

The Enrichment Program: Human and Physical Mediation

Unlike the posttest and TR sessions in which the students were just provided with preplanned computerized mediation from the CDRT (used in the posttest) and the CDRAT (used in the transcendence) and sometimes from the mediator, in the EP sessions other "mediational sources, i.e., human mediators (the researcher and the learners) and physical mediators (dictionaries) [either those available in the participants' smart phones or the ones in paper]" (Ableeva, 2010; p. 206) were utilized to help the participants better understand the texts. This means that the students could consult their peers, the mediator, or any other available physical tools to overcome the problems they faced while discussing the texts because the aim of all EP sessions in DA which were conducted in a group format was to overcome the encountered problems elicited from the pretest and DIALANG results.

Posttest

In the other stage of the design of this study, i.e., the posttest, two scores were obtained through taking the results of the CDRT test as follows: actual or NDA score (i.e., without mediation or the first try of the participants) and mediated (DA) scores. The actual score is what the students would have gotten in a traditional test where only the first responses counted, but the mediated score is the weighted score that includes points earned for attempts beyond the first.

Transcendence or Transfer Assessment Sessions

At last Transcendence as one of the main attributes of Feuerstein's framework in fully integrating assessment and instruction (Poehner, 2005) was taken into consideration and two Transfer Assessment (TR) sessions, namely TR1 ('near transfer' session) and TR2 ('far transfer' session) were held so as to reveal the extent to which learners were able to maintain their performance in novel and increasingly more challenging tasks. In other words, to assure the extent to which the participants could internalize and maintain the mediation provided in the course of previous sessions was the aim of the fourth stage of the design of this study. The importance of conducting TR sessions has been greatly underscored by Poehner (2005) who contended that in order to assure that the participants have really improved during DA sessions, transcendence should be taken

closely into consideration. That is, it is necessary to hold transfer assessments to claim that a change in the participants' performance has happened through DA procedures. During TR1 and TR2, the participants were provided with two authentic reading passages (two passages for TR1 and two passages for TR2) of higher difficulty in comparison with the ones which were covered in the pre- and posttest stages of the study. This means that the difficulty of the texts used in the pretest and posttest sessions was less than the difficulty of the texts used in TR1 and consequently TR2 had the most difficult texts in this study. The difficulty of the texts was measured by the research-based Free Text Readability Consensus Calculator.

Results

The posttest results were indicative of the individuals' development. What DA seeks for, however, goes far beyond development in a given situation and leads learners to apply what they have learned to novel and increasingly more challenging and unimagined situations in the future while assessing the same principles as the ones which were assessed in the pre- and post-test; a point which is the core purpose of transcendence (Poehner & Lantolf, 2013). According to Poehner and Lantolf (2010), TR reinforces individuals' internalization of knowledge and can result in their self-regulation when they attempt to. It is also, based on Poehner (2007, p.334), "fundamental to tracking development because it involves going beyond the test as learners endeavor to re-contextualize their abilities while engaging in new tasks." Several variables such as the length of sentences, complexity of words, the mood of examinees in the exam day, and so on affect individuals' performance (Cronbach, 1990); therefore, to ensure the individuals' extent of development while encountering novel and increasingly complex problems, the results obtained from CDRA1 and CDRA2, two software programs which were used to achieve the purpose of TR, along with the ones of mediated posttest (DA) are reported as follows.

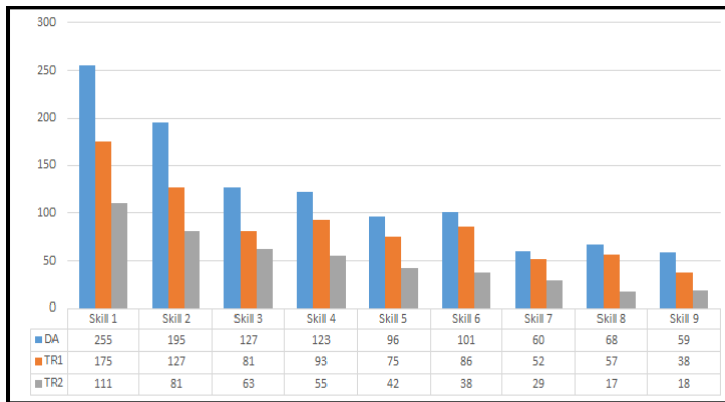


Figure 7. The Total Amount of Mediation Used in the Posttest (DA), TR1, and TR2 in Each Reading Skill

As the figure represents, not only did increased task complexity not lead to regression of students' development, but also it had an effect on increasing their development. In other words, the performance of no one of the at-risk investigated students deteriorated upon complication of the tasks in TR assessments. Of the transfer assessments, the participants found TR1 more challenging than TR2 since the amount of mediation they used within TR1 in all skills, with no exception, is more. This parallels Poehner's (2008) findings that the participants of his study reported more

challenges in TR1 in comparison to TR2; the numerical results were representative of their report as well. However, the results are in controversy with the study conducted by Ableeva (2010), in which the participants comprehended the near transfer (TR1) that was about a documentary on smoking in French restaurants better than the far transfer (TR2) which was a short (twenty-second) radio advertisement about a Belgian restaurant chain.

Transcendence which is included in tests incorporating mediation (Poehner & Lantolf, 2013) was utilized here to determine if the learners' development is sustained within increasingly complicated situations. The reasons upon inclusion of TR1 and TR2 were the feasibility of occurrence of microgenetic development even over a single exposure to learning (Wertsch, 1985) which encouraged the researchers to examine the participants' progress through TR1 and the necessity for the development to be ongoing (Feuerstein, et al. 1988) which in itself made TR2 be investigated as well so as to preclude interpretations of a haphazard improvement. With regard to microgenesis, it should be stated that though it was mostly explored in interactionist studies (e.g. Ableeva, 2010; Lantolf & Poehner, 2011; Poehner, 2008), it is, in the present study, investigated in an interventionist research to shed some more light on Poehner and Lantolf's (2013) findings. The results of the above table are representative of a microgenetic development in the DA part which is not only sustained but progressed within the two TR assessments as well.

Although research conducted on TR is indicative of a sharp difference (often failure) between individuals' performance in stages before TR and within TR (Brown & Ferrara, 1985), the individuals in this study, aligning with some of the ones in Poehner (2007 & 2008), both sustained and even improved their performance in spite of increased complexity of the passages. That is, in Poehner and Lantolf's (2013) words, the participants have internalized the mediation offered to them in earlier stages in such a manner that they required less amount and quality of mediation even in more complex tasks. Hence, requiring less mediation even after increasing the difficulty level of the passages, which assessed the same skills as the earlier ones, is clearly indicative of their internalized progress. Regarding the seventh skill, for instance, the participants required 60 hints to complete the two items of the given skill but abated them to 52 and 29 in TR1 and TR2, respectively.

Quality of the required mediation - either implicit or explicit- is another factor which in addition to its amount buttresses the numerical claim of the individuals' development of the study. Underscoring the importance of quality of mediation in some studies (e.g. Poehner, 2005 & 2007; Poehner & Lantolf, 2013) and its "change over time" (Poehner, 2005, p. 142) was a motive for the researchers to investigate it in the current study. In the same vein, Poehner and Lantolf (2013, p.327) more specifically stated that in TR in addition to tracking learners' development, it should also be "qualitatively [noted] whether and how much mediation they required to make these transitions." Owing to space constraints, only the first participant is taken into close consideration here to elucidate her performance with regard to the number of hints she used in each skill of DA, TR1, and TR2. Selecting this participant was done purposelessly. Through the performance of this participant in TR assessments, one can come to a fuller understanding of the extent the present learner has progressed or regressed and simultaneously paves the way for the researchers to ensure about her weaknesses and to predict how much mediation she would probably require in future tasks.

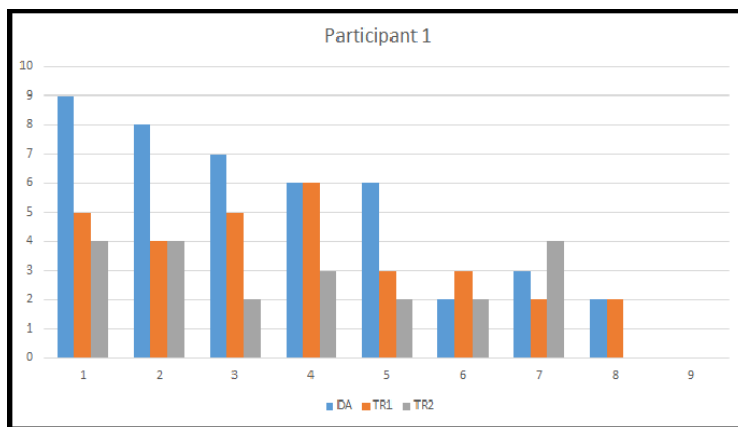


Figure 8. The Quality and Quantity of Mediation Used in the Posttest (DA), TR1, and TR2 in Each Reading Skill by Participant 1

Little or no; in case of skills 8 and 9 with regards to TR2 at least, requirement of mediation signifies both her independent performance and the degree to which she has internalized the intervention offered to her previously. Although TR tasks were more challenging than the ones encountered in DA and the learner was expected to require more intervention to work out the tasks, figure 8 depicts that she lessened the number of used hints in four of the skills (1, 2, 3, and 5), equaled in three ones (4, 8, and 9), and increased the use of hints in only two ones (6 and 7) unexpectedly. This is in controversy with the results obtained from the research on TR (e.g. Brown & Ferrara, 1985) indicating learners' drastic backshift or regression in case of extending their abilities. As noted, not in all skills was she able to progress. In other words, her problems reemerged in TR1 of skill 6 and TR2 of the 7th skill and she maintained her performance the same in the fourth skill along with the 8th and 9th ones. Therefore, by appraising her in TRs and failing to sustain her previous performance with regard to paraphrasing a specific sentence, it became apparent that some instruction is still required to bridge up the gaps related to paraphrasing through re-contextualizing the current and other underlying skills in more challenging tasks.

In a nutshell, the contentions made by generalizability are sharply doomed by the ones made by transcendence. Even though generalizability is mainly pertaining to the results obtained from a group of individuals to the remaining individuals who have some characteristics in common (Poehner, 2007), here the case is different in the sense that only the current-study participants' abilities are concerned through various stages of the study including pretest, posttest, TR1, and ultimately TR2. In other words, all of the participants would be simply regarded as mere at-risk as long as their pretest performance was taken into account but their manifest microgenetic development in DA (posttest) and its sustenance afterwards in TRs even under engaging in more demanding re-contextualized tasks sealed the contrast between what generalizability and TR seek for. Their TR performance, furthermore, approves the title of Feuerstein et al.'s (1988) book: *Don't accept me as I am: Helping "retarded" people to excel* which, in turn, shows the need to DA reassessments to help at-risk learners excel. This, in addition, indicates that there is no endpoint to progress and merely gaining a low score should not be interpreted as the inability of a special learner to surpass.

Discussion and Conclusion

Transcendence was investigated by putting the individuals into two novel and increasingly more challenging and unimagined situations (TR1 and TR2) in which the same qualities were questioned anew (Poehner & Lantolf, 2013). Regarding TR, Feuerstein et al. (1988) presented the term 'active modification' which seems to be in parallel with what TR sets for, i.e. helping individuals develop their abilities to be capable of managing more various contexts. Ibid. (p.14) stated that 'active modification' is interested in "increasing the individual's modifiability and enhancing his adaptational capacities." The results, as the ones in Poehner et al.'s (2014) study, were indicative of not only their performance maintenance in the posttest but also their progression in terms of some reading skills. The individuals' sustainability and progression, in some skills, revealed the concomitance of the results with the ones in Poehner's (2007 & 2008) studies. In conclusion, regarding the obtained results, it was disclosed that there is no endpoint to progress and merely gaining a low or high score should not be interpreted as the inability of a special learner to surpass.

The study, ultimately, dealt with counterarguments regarding the standard psychometric dimensions of the obtained findings of the study, i.e. reliability and validity. Unlike NDA advocates who are mainly concerned with the impact of any sort of instruction or mediation on modification of learners' improved performance and believe in the stability of individuals' abilities (Sternberg & Grigorenko, 2002a), DA proponents being forerun by Vygotsky, on the other hand, underscore that "development potentially never stops and as long as humans have access to new symbolic mediational means through social interactions (e.g. teaching/learning processes), they continue to develop." (Ableeva, 2010, p.143)

With regard to reliability, NDA supporters advocate that offering mediation during assessment leads to unreliability because it influences learning. Poehner's (2008) reasoning, in reaction to this argument, seems to be satisfactorily compelling. He contends that this DA - NDA controversy emanates from the fact that NDA views learners' abilities as stable points and by so doing it has room for claiming reliability, whereas DA believes in modifiability and amenability of learners' abilities to changes, it also discards the assessor's indifferent or disinterested position throughout the testing procedure and ultimately motivates the assessor, in Lantolf and Poehner's (2004, p.59) words, to "react to learners' responsiveness." Owing to the modifiability of learners' abilities and cognitive functions, hence, reliability which is a pivotal psychometric property of tests is irrelevant to the DA goals and should be casted some doubt on.

Between the two psychometric terms, it is validity which, according to some scholars (e.g. Lantolf, 2006; Lantolf & Poehner, 2008)) seems to make more sense in the context of DA. This concept is more relevant since any assessment is considered valid provided that it is advantageous and beneficial to learners. In the same vein, Lantolf and Poehner while underscoring the amenability of the future to change (indicating the irrelevance of reliability to DA goals) noted that one can gain a better understanding of the validity of a DA procedure through the extent of development made within the procedure.

The most important significance of this study is its transcendence of learning beyond the assessment context by holding two transfer assessment sessions (TR1 and TR2). In this study, the researchers sought to find out whether the participants could extend their development, if any, into more challenging and novel assessment contexts. If the number of mediator's prompts or hints decreases during the DA sessions and it continues to fall more in TR1 and even more in TR2 in a way that during TR2 no mediation is offered at all by the mediator, then it can be stated that the participants have gained greater control of their performance and are able to function without intervention from the mediator. Although the learners learn in DA, they forget to maintain their learning and we have backsliding which is also considered as development by Vygotsky.

It is noteworthy as well to mention the two common formats of DA and to state which format and how they were utilized to reach the aims of the study. Owing to using the C-DA procedures, the present study follows the interventionist approach to DA; thus, it utilizes the sandwich format because of its test-teach-test paradigm too (Sternberg & Grigorenko, 2002), but it also makes use of cake format because of its item-by-item instruction (*ibid.*) and also because of the CDRT and CDRAT software programs which are developed in a way that provide hints or mediation based on the examinee's response to each test item. In summary, the study uses the interventionist approach to DA and a combination of the two common formats of DA, i.e. sandwich and cake.

The study also had some limitations which need to be explored further in the future. First, a shortcoming of the EP sessions was that the individuals' general problem areas were considered to be of importance and notice. Though it is appropriate to "examine an individual's C-DA scores collectively rather than in isolation" (Poehner et al., 2014, p. 13), it is also important to take his/her scores together to conclude the most nuanced and comprehensive information on the specific type of support he/she needs to resolve the problem areas for different reading comprehension constructs. Second, though it was attempted to reduce the possibility of guessing throughout the TRs, the researchers failed to recognize or determine when and in what items they were more probable to have guessed. In the present study, the participants were offered a further explanation upon completion of any item correctly. So it was not visible if the participants were eager to read the explanation or not. Ultimately, the study narrowed the scope of its investigation to just one skill; reading. Though it broadened its horizon to a greater extent in comparison to the ones done earlier (Pishghadam et al., 2011; Shabani, 2012; Teo, 2012, 2014), carrying out research on the foregoing limitations might shed some more light on the area under question, i.e. reading comprehension. Other skills (speaking, listening, and writing) and even language subskills such as grammar and vocabulary can further explore the effectiveness of DA, in general, and C-DA and other classroom-based procedures to DA (e.g. Group Dynamic Assessment or G-DA in short), in particular.

References

- Aljaafreh, A., & Lantolf, J.P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 78, 465-483.
- Ableeva, R. (2010). *Dynamic Assessment of listening comprehension in second language learning* (Doctoral dissertation, Pennsylvania State University).
- Azabdaftari, B. (2013 a). An explication of concordance between man's mental structure and the narrative structure in the light of Vygotsky's SCT. *Iranian Journal of Language Teaching Research*, 1(3), 45-51.
- Azabdaftari, B. (2013 b). On the implications of Vygotskian concepts for second language teaching. *Iranian Journal of Language Teaching Research*, 1(2), 99-114.
- Birjandi, P., & Ebadi, S. (2012). Microgenesis in dynamic assessment of L2 learners' socio-cognitive development via web 2.0. *Procedia-Social and Behavioral Sciences*, 32, 34-39.
- Cronbach, L. J. (1990). *Essentials of Psychological Testing* (5th ed.). New York: Harper & Row.

- Ebadi, S. (2014). L2 Private Speech in Online Dynamic Assessment: A Sociocultural Perspective. *Iranian Journal of Applied Linguistics*, 17(1), 49-70.
- Farhadi, H., Jafarpur, A., & Birjandi, P. (1994). *Testing language skills: From theory to practice*. Tehran: SAMT.
- Farr, J. N., Jenkins, J. J., & Paterson, D. G. (1951). Simplification of Flesch Reading Ease Formula. *Journal of applied psychology*, 35(5), 333- 337.
- Feuerstein, R., Rand, Y., & Rynders, J. E. (1988). *Don't Accept Me as I Am. Helping Retarded Performers Excel*. New York: Plenum.
- Hernández-Murillo, R., & Shell, H. (2014). The Rising Complexity of the FOMC Statement. *Economic Synopses*, 23. Available at: <http://research.stlouisfed.org/publications/es/article/10252>.
- Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL text comprehension of at-risk students. *School Psychology International*, 23(1), 112-127.
- Lantolf, J.P. (Ed.). (2000). *Sociocultural Theory and Second Language Learning*. Oxford: Oxford University Press.
- Lantolf, J. P. (2006). Sociocultural theory and L2: State of the art. *Studies in Second Language Acquisition*, 28(1), 67-109.
- Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment of L2 development: bringing the past into the future. *Journal of Applied Linguistics*, 1(2), 49-72.
- Lantolf, J. P., & Poehner, M. E. (2011). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research*, 15(1), 11-33.
- Lantolf, J. P., & Poehner, M. E. (2013). The unfairness of equal treatment: objectivity in L2 testing and dynamic assessment. *Educational Research and Evaluation*, 19(2-3), 141-157.
- Lidz, C. S., & Gindis, B. (2003). Dynamic assessment of the evolving cognitive functions in children. In *Vygotsky's Educational Theory in Cultural Context*. A. Kozulin, B. Gindis, V. S. Ageyev, S. M. Miller (Eds.). Cambridge: Cambridge University Press.
- Lidz, C. S., & Macrine, S. L. (2002). An alternative approach to the identification of gifted culturally and linguistically diverse learners, the contribution of dynamic assessment. *Mensa Research Journal*, 33, 15-35.
- Linn, R. L. (2000). Assessments and accountability. *Educational researcher*, 29(2), 4-16.
- Lynch, B. K. (2001). Rethinking assessment from a critical perspective. *Language Testing*, 18(4), 351-372.
- McNamara, T. (2001). Language assessment as social practice: Challenges for research. *Language Testing*, 18(4), 333-349.

- Moss, P. A. (1996). Enlarging the dialogue in educational measurement: Voices from interpretive research traditions. *Educational Researcher*, 25(1), 20-29.
- Oskoz, A. (2005). Students' dynamic assessment via online chat. *CALICO Journal*, 22(3), 513-536.
- Pichert, J. W., & Elam, P. (1985). Readability formulas may mislead you. *Patient Education and Counseling*, 7(2), 181-191.
- Pishghadam, R., & Barabadi, E. (2012). Constructing and Validating Computerized Assessment of L2 Reading Comprehension. *Iranian Journal of Applied Linguistics*, 15(1), 73-95.
- Pishghadam, R., Barabadi, E., & Kamrood, A. M. (2011). The Differing Effect of Computerized Dynamic Assessment of L2 Reading Comprehension on High and Low Achievers. *Journal of Language Teaching & Research*, 2(6), 1353-1358.
- Poehner, M. E. (2011). Dynamic Assessment: Fairness through the prism of mediation. *Assessment in Education: Principles, Policy & Practice*, 18(2), 99-112.
- Poehner, M. E. (2008). *Dynamic Assessment: A Vygotskian Approach to Understanding and promoting L2 development*. USA. Springer Science.
- Poehner, M. E. (2007). Beyond the test: L2 dynamic assessment and the transcendence of mediated learning. *The Modern Language Journal*, 91(3), 323-340.
- Poehner, M. E. (2005). Dynamic assessment of oral proficiency among advanced L2 learners of French. Unpublished dissertation. Pennsylvania State University, University Park.
- Poehner, M. E., Zhang, J., & Lu, X. (2014). Computerized dynamic assessment (C-DA): Diagnosing L2 development according to learner responsiveness to mediation. *Language Testing*, Special Issue, 1-21.
- Poehner, M. E., & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during Computerized Dynamic Assessment (C-DA). *Language Teaching Research*, 17(3), 1-21.
- Poehner, M. E., & Lantolf, J. P. (2010). Vygotsky's teaching-assessment dialectic and L2 education: The case for dynamic assessment. *Mind, Culture, and Activity*, 17(4), 312-330.
- Shabani, K. (2012). Dynamic assessment of L2 learners' reading comprehension processes: A Vygotskian perspective. *Procedia-Social and Behavioral Sciences*, 32, 321-328.
- Shohamy, E. G. (2001). *The power of tests: A critical perspective on the uses of language tests*. Pearson Education.
- Shohamy, E. L. A. N. A. (1998). Inside the "black box" of classroom language tests. *Studia Anglica Posnaniensia*, 33, 343-352.
- Teo, A. (2012). Promoting EFL Students' Inferential Reading Skills through Computerized Dynamic Assessment. *Language Learning & Technology*, 16(3), 10-20.

- Teo, A. (2014). Beyond Traditional Testing: Exploring the Use of Computerized Dynamic Assessment to Improve EFL Learners' Reading. *Arab World English Journal*, 5(1), 42-58.
- Tzuriel, D., & Shamir, A. (2002). The effects of mediation in computer assisted dynamic assessment. *Journal of Computer Assisted Learning*, 18, 21–32.
- Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. MA: Harvard University.
- Wang, L. W., Miller, M. J., Schmitt, M. R., & Wen, F. K. (2013). Assessing readability formula differences with written health information materials: application, results, and recommendations. *Research in Social and Administrative Pharmacy*, 9(5), 503-516.
- Wertsch, J.V. (1985). *Vygotsky and the Social Formation of Mind*. Cambridge, MA: Harvard University Press.

Saman Ebadi is currently working as an assistant professor in the department of applied linguistics at Razi University, Kermanshah, Iran. His areas of interest are dynamic assessment, sociocultural theory, qualitative research methodologies, ELT material development and discourse analysis. He has presented and published papers in international conferences and journals.

Abdulbaset Saeedian holds an MA in Applied Linguistics at Razi University, Kermanshah, Iran. His areas of interest are dynamic assessment, discourse analysis, and attitudes towards language learning. He has presented and published papers in international conferences and journals.