PLANNING FUTURE INSTRUCTIONAL PROGRAMS THROUGH COMPUTERIZED L2 DYNAMIC ASSESSMENT

by **Saman Ebadi** and **Abdulbaset Saeedian**

Razi University

P. O. Box 67149-67346, Baghe Abrisham, Kermanshah, Iran

samanebadi @ gmail.com, saeedianbaset @ gmail.com

Abstract

Dynamic Assessment (DA) is a postmodern notion in testing which sees instruction and assessment as inextricably mingled contending that learners will progress if provided with dynamic interactions. The main purpose of the study is to see if the scores generated by the computerized dynamic assessment (C-DA) which is grounded in Vygotsky's theoretical framework in congruence with the concept of DA can lead to designing a syllabus which results in the participants' reading comprehension development. In the present study, a total of 32 Iranian EFL undergraduates from a university in Iran were selected on the availability basis. The study made use of the interventionist approach (the same mediation for all individual learners) to DA due to a two-fold aim: being more economically-supported and owing to its feasibility in focusing on larger cohorts of individuals. Investigating the learners' generated scoring profiles through CDA revealed that not only did the learners have varying problem areas but also they needed different amount of mediation for identical test items. These profiles reiterated the fact that learners with different zones of proximal development (ZPDs) require customized instructional programs to reflect their individualized needs.

Keywords*:* instructional program, dynamic assessment, computerized dynamic assessment, interventionist DA

1. Introduction

Dynamic Assessment (DA) is an emergenistic and postmodern notion in testing (Pishghadam & Barabadi, 2012) which sees instruction and assessment as inextricably mingled and not as separate processes (Haywood & Lidz, 2007; Haywood & Tzuriel, 2002; Lantolf, 2009; Lidz & Gindis, 2003). It is based on dynamic interaction between the examiner and the examinee (Birjandi & Ebadi, 2012), in which the former helps the latter achieve their best. DA, which is rooted in mediated teacher-learner interactions, has some advantages, including providing deeper insights into how individuals' abilities change and develop over time (Ableeva, 2010). Mardani and Tavakoli (2011, p. 695) remarked that another advantage of DA is its fairness,

stating that "DA is an integral part of the assessment but not its entirety, because no one approach can provide adequate answer to all questions."

In general, there are two approaches to DA:

1. in the interventionist approach to DA, the same mediation is used with every learner, therefore, it is easier to manage a larger number of participants (Poehner, 2008).

2. in the interactionist approach to DA, the mediator cooperates separately with each learner to co-construct ZPDs during different one-on-one sessions and the mediation provided for each student may be (is) different from the one provided for the others (Aljaafreh & Lantolf, 1994). In other words, as Poehner (2008, p. 20) stated, "proponents of interactionist DA follow a case study approach to research and validate their work on the basis of an accumulation of in-depth studies of individuals or groups of individuals."

This study utilized DA to explore EFL learners' reading comprehension which is "the process of simultaneously extracting and constructing meaning through interaction and involvement with written language" (Snow, 2002, p.11). As students advance in school, researchers suggest reading instruction should become more disciplinary, reinforcing and supporting students' academic performance (Shanahan & Shanahan, 2008). Based on the results of a pre-test and in line with the students' needs or areas of problem, Beck, McKeown, and Kucan (2013) provided some reading strategies such as learning to identify and state the main idea by naming the who or what (the main person, animal, place, or thing the selection is about), telling the most important thing about the who or what, etc. Having analyzed the pretest results, Beck et al. (2013) provided the learners with reading comprehension strategies developed to meet the considerable instructional needs of the at-risk students participating in the study.

Since it is a challenging and unmanageable task for many EFL teachers to provide one-to-one mediation to individual students (Teo, 2012), computer software called Computerized Dynamic Reading Test (CDRT) was originally developed by Pishghadam and Barabadi (2012), which offered the learners pre-fabricated and standardized mediation in appropriate time, i.e., whenever it was requested by the learners. The software also provided the researchers with the learners' scoring file consisting of a DA score, a non-dynamic assessment (NDA) score, the total number of mediation used by each individual, and the amount of time spent on completing the test. The software developers sought to overcome the time-constraint challenge, which is one of the major problems many EFL teachers are

struggling with. Thus, technology was utilized in the study to check the role of the computerized dynamic assessment (C-DA) principles in unifying teaching and assessing in general in future teaching programs using the interventionist DA. Poehner (2008, p. 43) advised that in studies with large cohorts of participants the interventionist DA be used. Due to a relatively large number of participants in this study, the interventionist DA was employed because it "is more in line with Vygotsky's vision of how the ZPD can be used to reorient education to learner development and is therefore more relevant to the classroom."

This section has briefly sketched how teaching and learning have been modified in the digital age and how teachers need to take into due account problems that learners may encounter when reading and writing digital texts in English. Below, some possible barriers in digital literacy will be identified to propose possible solutions in terms of teaching and learning strategies. Next, the development of the needs analysis will be explained, followed by indicating the area of interest for Group 3 (learners of 15-18 age range), i.e. New Travel. Finally, the construction of a web corpus to develop a web browser from the specifications found through empirical data will be described, which emerged during the first stage of project. Some conclusions will be presented, indicating future developments.

2. Literature review

2.1. Dynamic Assessment

DA, which is a pedagogical approach and a development-based activity (Poehner, 2005), is theoretically framed within the works of Vygotsky and contends that, unlike traditional testing methods, instruction and assessment are dialectically integrated. Some key concepts lie at the heart of the notion of DA. The first important notion is mediation; the process by which other-regulated activities are transformed into self-regulated ones (Lantolf & Thorne, 2006). In the same vein, Aljaafreh and Lantolf (1994) proposed a model of mediation from other-regulation to self-regulation in learners, which included five transitional levels starting from the most implicit or indirect to the most explicit or direct. These levels are, in fact, indicative of three stages: object-regulation; other-regulation; and self-regulation. As the theoretical underpinnings of DA, mediation and regulation are of great importance to understand which type of mediation should be offered to whom, at what extent, and when. Practically, this is a tremendous task to do in educational contexts in cases where the interactionist DA is utilized, as "the levels [are] not determined in advance of the study" (Aljaafreh & Lantolf, 1994, p. 471).

According to Lantolf (2000), one of the forms of mediation is regulation. Frawley and Lantolf (1985) defined regulation as the way in which an individual sees a task and also their ability to successfully complete that task. It is one form of mediation that goes through three stages to complete its process. The stages, respectively, are as follows:

In the first stage of object-regulation, individuals use objects in their environment in order to think. That is, an object tells us to do something; a persuasive advertisement, for instance. In this regard, Poehner (2008, p.27) commented that "[a]t the level of object regulation, psychological functioning is controlled by the environment rather than by the individual, and so in response to hunger the individual eats what is immediately available or goes in search of food."

In the second stage of other-regulation, individuals' performance is primarily controlled by someone else (Lantolf & Poehner, 2011). That is, it includes implicit and explicit mediation by parents, peers, teachers, so forth. Here someone tells us to do something; for instance, a mother tells her child to do his/her homework.

In the third stage of self-regulation, minimal or no external assistance is required from the individuals' side to accomplish activities. In other words, individuals establish control over their own performance (Lantolf & Poehner, 2011). We tell ourselves to do something; for instance 'I need to finish my M.A. thesis before Ramadan.' In fact, self-regulation enables us to control our responses in order not to merely act instinctively but instead choose from among possible alternatives intentionally (Poehner, 2008). Preferring not to eat anything in an effort to lose weight while being invited by a friend of yours is an example of this kind.

In this regard, Vygotsky (1978) argued that moving from other- or object- regulation to self-regulation is the primary way in which humans develop higher-order thinking skills. In other words, a learner has to pass from being object-regulated to being self-regulated for development to occur (Summers, 2008). This movement is termed 'Internalization', a process through which higher mental functions are created.

The importance of the type of mediation or interaction which is provided for learners is reflected in Vygotsky's beliefs, who stated that learning occurs as the result of interaction, but not any kind of interaction, i.e. it only emerges as the result of interaction within the ZPD. The theoretical underpinning of DA (Kozulin & Gindis, 2007) implies that potential development differs from actual development (Poehner & Lantolf, 2005). That is to say, what the individual is able to do one day with assistance s/he is able to do tomorrow alone. This means that depending on an individual's ZPD, the mediator should match the provided interaction to that person's potential for better results. Out of what has just been stated, it can

be understood that people's ZPD is not fixed but instead it is a malleable and open-ended trait of them, which can become apparent through interaction and consequently develop the potential for learning, of course, if suitable opportunities are provided (Wells, 1998).

As it is clear from this discussion, not all ZPD-based studies can be conducted without relevant help or assistance. Any assistance from the mediator's side should have two important properties (mechanisms) to be effective: First, it should be gradual, second, it should be contingent. Different researchers have used different terms to refer to these two properties. Summers (2008) referred to these mechanisms as 'quality mediation', for instance. Any help which has these two properties is referred to as 'ZPD-based help' (Tajeddin & Tayebipour, 2012) or 'negotiated help' (Nassaji & Swain, 2000). If it does not have these mechanisms, it is called, according to the just-mentioned studies, 'random help'; that is there is no attempt to adjust the level of assistance to the learner's responsiveness. Aljaafreh and Lantolf (1994) defined the former, i.e., graduation, as help which moves from highly implicit level through more and more concrete levels until the appropriate level is reached. Of course, the assistance from the mediator's side should not be too explicit to let him/her take over more of the activity than is necessary. They also defined the latter, i.e., contingency, as help which "should be offered only when it is needed, and withdrawn as soon as the novice shows signs of self-control and ability to function independently" (p. 468). In another definition, Gibbons (2003, p. 267) stated that contingency consists of the "assistance required by the learner on the basis of moment-to-moment understanding." Tajeddin and Tayebipour (2012) called these two mechanisms as the building blocks of DA and claimed that many academic disciplines have utilized them.

2.2. Dynamic Assessment vs. Dynamic Testing

The difference between dynamic assessment and dynamic testing is reflected in Sternberg and Grigorenko (2002), who remarked that "[i]n essence the goal of dynamic assessment is to intervene and to change. The goal of dynamic testing, however, is much more modest - it is to see whether and how the participant will change if an opportunity is provided" (p. 30). According to Sternberg and Grigorenko (2002), dynamic testing occurs if two components, i.e., assessment and pedagogical intervention, are combined. Therefore, it can be concluded that dynamic testing provides prefabricated mediation for students to find out how much they will or will not change when offered pre-determined assistance.

Although Sternberg and Grigorenko were determined to highlight the differences between these two terms, "dynamic assessment and dynamic testing should not be thought of

as separate enterprises" (Poehner, 2008, p. 17). By the same token, and without considering the differences between these two terms, having chosen Vygotsky's discussion of microgenesis which deals "with the issue of development occurring very quickly (Poehner, 2008, p. 18)", the present researchers adopted C-DA to be used throughout this study to refer to those sessions which aim at unifying assessment–instruction as the basis of the DA procedures.

2.3. Computerized Dynamic Assessment (C-DA)

In congruence with the concept of DA, the computerized dynamic assessment (C-DA) is grounded in Vygotsky's theoretical framework (1978). Some studies have been conducted in the field of education on C-DA.

For instance, Tzuriel and Shamir (2002) conducted a study in the area of C-DA and tailored mediation to learners. They attempted to assess kindergarten children's seriational thinking abilities because they believed that these abilities were central to success in learning mathematics. The prompts have been prefabricated and arranged from implicit ('try again') to explicit (providing more relevant information about the item in question). As it is clear, it follows an interventionist approach to DA because the prompts are prefabricated but since teachers are also allowed to take part in the administration of the test actively, i.e. provide supplemental support for learners who fail to answer the questions correctly, just like interactionist DA. The authors stated that more in-depth diagnoses of learner abilities is provided through this procedure when teachers are present in comparison to the time when the mediation is only provided by computer.

Another study conducted within this domain was the one by Pishghadam and Barabadi (2012). Underscoring the increasing importance of DA in second language and reading comprehension, the researchers magnified the shortcomings of DA and paved their own way for introducing their own developed software called CDRT to examine L2 reading comprehension through C-DA. To justify what they have done, they cited some interactionist studies which based on the authors followed a sandwich format, though such a claim cannot be supported based on the seminal work done by Sternberg and Grigorenko (2002), Poehner or Lantolf. They claimed that the problem of interactionist studies is that the number of their participants is low, while in sandwich format studies the mediation phase and the assessment session are administered separately from each other. In other words, instruction and assessment are not fully integrated in interactionist studies which follow the sandwich format.

They also contended that in addition to these shortcomings, interactionist DA does not take the psychometric properties of testing into consideration.

To solve these problems, Pishghadam and Barabadi (2012) used C-DA which is interventionist and follows the cake format. Though not an unbreakable principle, the general consensus is that the interventionist studies tend to follow the sandwich format because of their assessment-instruction-assessment type. However, the reason why Pishghadam and Barabadi (2012) claimed C-DA, though being interventionist, follows the cake format is the mediation which the CDRT software provides for learners whenever problems arise during the administration of the assessment. Their study can also be regarded as a study in which sandwich format has been used because in addition to the mediation provided in the pre-test for any individual items of the reading comprehension questions, mediation was also provided to students based on their pre-test performance. This mediation was provided for students in a separate way from assessment. That is, while having no assessment session, the students were mediated to be more prepared for the post-test. The following advantages of C-DA were mentioned in Pishghadam and Barabadi (2012, p. 79) as well: "1) reliability and validity are taken into account; 2) many students can be assessed dynamically, and 3) mediation is given at the time of assessment not in a separate session."

The two most prominent figures of DA, i.e., Poehner and Lantolf, carried out a study on the domain of C-DA in 2013 to show its application to larger classes. Focusing on the significance of the instructional quality of mediation, they referred to a phenomenon called 'microgenesis', which Wertsch (1985) considered as a process that provided opportunities for development simultaneously even during a single session. While microgenesis primarily deals with a context in which learners and mediators have a moment-to-moment interaction (Lantolf & Poehner, 2011), their study explored the principles of mediation into a computerized approach to DA. In fact, their study explored microgenesis in the context of C-DA taking L2 Chinese, French, and Russian listening and reading comprehension into consideration. They designed some tests and aimed to differentiate between the learners' independent and mediated performance, to foresee the difference between their mediated and non-mediated performance (learning potential), and finally to reassure evidence of learning by applying the concept of transcendence into the tests. Similarly to Poehner (2005), the number of semesters the participants had spent studying in university (here intermediate level because they had studied four semesters) was taken as a way to determine the participants' proficiency level. In that study two skills (reading comprehension and listening comprehension) were taken into account for the learners of two languages (Chinese and French), with the gain score

or Learning Potential Score (LPS) for any one of these skills and also the reliability coefficient of the tests for the aforementioned skills calculated.

As it is clear, C-DA has several advantages including simultaneous administration to large numbers of learners; providing learners with the opportunity to reassess as many times as they would like; and informing the test takers of their performance in the test automatically after they respond the exam. This, however, does not mean that C-DA is flawless. Though it overcomes some of the shortcomings of other approaches to DA, it faces the same major challenge as all other interventionist approaches such as Group Dynamic Assessment (G-DA): we cannot claim and know how learners' performance would differ if they were provided with other forms of mediation.

3. Methodology

In concomitant with most DA studies (Ableeva, 2010; Lantolf & Poehner, 2013; Poehner, 2005; Teo, 2012), this research also uses qualitative methodology which best fits DA principles (Ableeva, 2010) but it can be regarded as quantitative as well since it follows the interventionist approach to DA (Poehner & Lantolf, 2005). In other words, both qualitative and quantitative research procedures have been used in the study.

This study was guided by the following question: How useful are the scores generated by the computerized dynamic assessment to planning future teaching programs?

3.1. Participants

The participants of the study were drawn from all undergraduates of B.A. Teaching English as a Foreign Language from a university in Iran. From among the 47 available undergraduate students, 32 were non-randomly selected to take part in the study. The participants' age ranged from 22 to 31 years indicating the participants were adults, and English was the second language of these adult learners. The homogeneity of the participants was taken for granted by claiming this statement (also being contended by Poehner, 2005) that the number of semesters the students have spent studying a language shows the proficiency level of whom in that language. Of course, the results obtained from the DIALANG, a free online assessment system to determine learners' proficiency level, were also indicative of the homogeneity of the participants. 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.

The importance of using this study lies in the contradiction between the terms 'advanced', and 'at-risk' learners. Since the participants were seniors, they were considered as 'advanced' students but due to their low proficiency, based on the results obtained from the Placement Test of DIALANG, they were called 'at-risk' too. Therefore, it is of really great importance to reiterate that the tests which have been used in this study were all suitable for 'advanced' level students and that using DIALANG was just to reassure that students were 'at-risk'.

3.2. Instruments

3.2.1. The researchers as tools

It is not possible to separate the researcher from the research in qualitative studies (Merriam, 1997; Summers, 2008). This means that the researcher's impressions and perceptions of events influence data analysis. Due to the social nature of human beings and in line with the Vygotskian perspective, the researcher's role in carrying out qualitative research is demanding. To underscore the inseparable role of researchers in research, Smagorinsky (1995) stated that in the data collection procedure the relationship between researchers, participants, context of the study, and the means of data collection is of high importance. Hence, if a researcher contends to separate qualitative research into the area of SCT from the social situation, it can be stated that researcher has misinterpreted the Vygotskian cognitive theory (Summers, 2008).

Our position is that learning is a socially constructed event and it is thus reflected in the way we teach and assess learners. For us, the environment in which learning occurs is the actual source of learning and that it is not possible to consider learning, instruction and assessment as inseparable. This means that we were participant-observers who held the Enrichment Program (EP) sessions in DA and were actually the facilitators of the C-DA procedure. We also played another role as technology troubleshooter. Therefore, it can be claimed that we were a data collection tool and our presence affected the participants and the data collection. It is noteworthy that the whole data collection was done by the pre-test, the Computerized Dynamic Reading Test (CDRT) developed by Pishghadam and Barabadi (2012) in the post-test, and some Enrichment Program (EP) sessions in between.

3.3. Computerized Dynamic Reading Test (CDRT)

To see whether C-DA could assist the learners realize their learning potential or not, the researchers utilized the previously validated and reliable software developed by Pishghadam

and Barabadi (2012), namely Computerized Dynamic Reading Test (CDRT). With regard to the software, it is worth mentioning that it can easily run on any PC provided that the NET Framework software is installed on it. Students have to enter some information such as their name, age and major (students can choose a pseudonym to remain anonymous for other people but they should say it to the mediator) and after reading the software description go directly into the passage and answer the items while consulting the preplanned hints which are automatically shown if a wrong response is chosen. It takes about two hours to complete the test and after completing it a scoring file is created on the desktop to know about the test taker's performance.

3.4. Procedure

Regarding the design of the study, the following stages were monitored: the pre-test; the Enrichment Program; and the post-test. The first stage, i.e., the pre-test, consisted of two passages which were similar to the texts used in the DIALANG with regard to the degree of difficulty and included items which assessed the same areas the participants showed to have problems with (e.g. their inability to connect the ideas in the passages, their difficulty at identifying the main ideas of texts, etc.).

Having collected the pre-test results and consequently having identified the participants' problematic areas, the researchers determined the number of sessions to be held for the (Enrichment Program) EP (two weeks: two sessions per week; each session one and a half hours).

In the last stage of the design of this study, i.e., the post-test, 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. This means that the CDRT which was developed by Pishghadam and Barabadi (2012) was used in the post-test design of this study. Similar to the pre-test, a one-week period was determined to collect the data in this stage too because there were only seven computers available and the participants could not wait there for others to fulfill their job. In this stage which was done individually the students' score gained with the use of hints was termed 'dynamic' score and their score gained with no hint (i.e., their first try) was called 'non-dynamic' score.

In the pre-test, a total of 20 items each worth 5 points were included in the passages in accordance with the areas being questioned in the items of CDRT. It took one week to collect the data in this stage because the tests were in the paper-based form and there was no spacious class for 32 students at our language institute. Before the pre-test stage in which learners'

problems were identified, the EP (EP in DA) comprised an unknown number of instruction sessions and even the time which had to be dedicated to each task was not predictable in advance. This was also underscored by Nassaji and Swain (2000, p. 48), who claimed as follows: "Although it is preferable from research point of view to have equal time-on-task in experimental designs, the nature and the amount of negotiation required in the ZPD condition to complete the tasks collaboratively and successfully was quite unpredictable: it could not be fixed in advance." Hence, it is the participants' pre-test results that can determine the nature and quantity of interaction, not the mediator's intention.

Finally, the post-test stage followed the Enrichment Program. However, on the contrary to the EP, other "mediational sources" such as especially dictionaries were not allowed to be used so that it would be necessary for students to rely for word meaning on strategies such as prediction and hypothesis (Kozulin & Garb, 2002) which were instructed in the EP sessions. It is worth noting that in contrast to actual scores in the pre-test and post-test whose aims were to evaluate the participants' actual level of text comprehension, the purpose of mediated (DA) scores in the post-test was to evaluate the potential level of the students' L2 reading comprehension.

Upon completion of the test in CDRT, the learners were presented with two scores (DA and NDA) and the amount of mediation used for answering the test in a specific span of time. Therefore, to answer the study's research question and identify the more specific and nuanced impacts of the roles of scores generated by C-DA on planning a future teaching program, the participants' scores in nine reading skills were closely taken into account. The usefulness of scores to planning a teaching program which is considered as "an important question" by Lantolf and Poehner was proposed here to see if these scores can lead to designing a syllabus which results in the participants' reading comprehension improvement (Poehner, Zhang & Lu, 2015, p. 346).

Though each participant's scoring profile generated by C-DA was worth investigating, it was not practically possible due to a high number of skills and participants. Thus, since they all yielded high LPSs and due to limitations of space, 6 participants were selected purposively due to their distinguishing actual and mediated scores; 3 from the first 16 and 3 from the second 16 participants, to be explicated. Participants 1, 7, and 16 were selected from the first group, while participants 20, 22, and 26 from the second one. However, they were compared in the following pair: 1 and 22; 16 and 26; and finally 7 and 20.

Participants 1 and 22 were compared with each other because they produced the same actual and mediated scores and hence gain scores and LPSs. One may think that they require

the same amount of mediation or that they have problematic language areas in common simply because they have the same scores or performance.
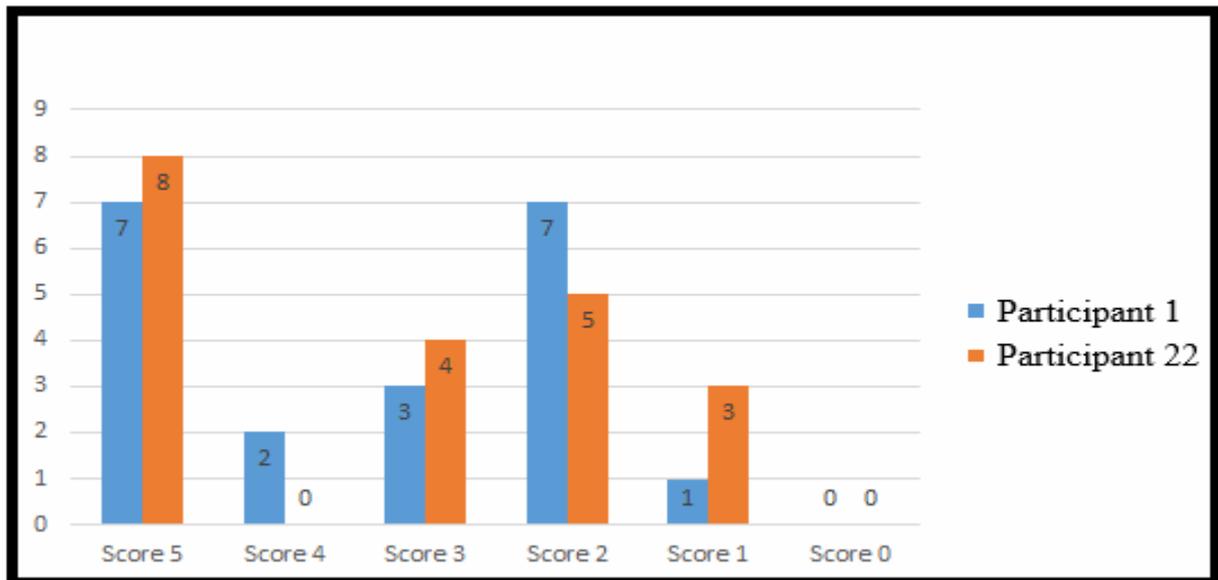

Figure 1. Different levels of mediation required for items by the participants

Investigating their generated scoring profiles unfolded that not only did they have varying problem areas but also they needed different amount of mediation for identical test items. For instance, the Figure shows that Participant 1 responded to 7 items without any mediation but required 1 hint for 2 items, 2 hints for 3 items, 3 hints for 7 items, and 4 hints for 1 item while Participant 22 answered 8 items without any mediation and though he did not score 4 in any items, he required 2 hints for 4 items, 3 hints for 5 items and ultimately 4 hints for 3 items. As it is illustrated in Figure 1, none of the two required 5 hints for any one of the items; meaning that they were able to answer the items before the answer was shown on the screen. As in Poehner et al. (2015), the results of this study showed that simply producing identical actual, mediated, etc. scores does not mean that learners need the same amount of assistance as well. To make sure about their strength or weaknesses in the nine reading comprehension skills, Figure 2 should be consulted.
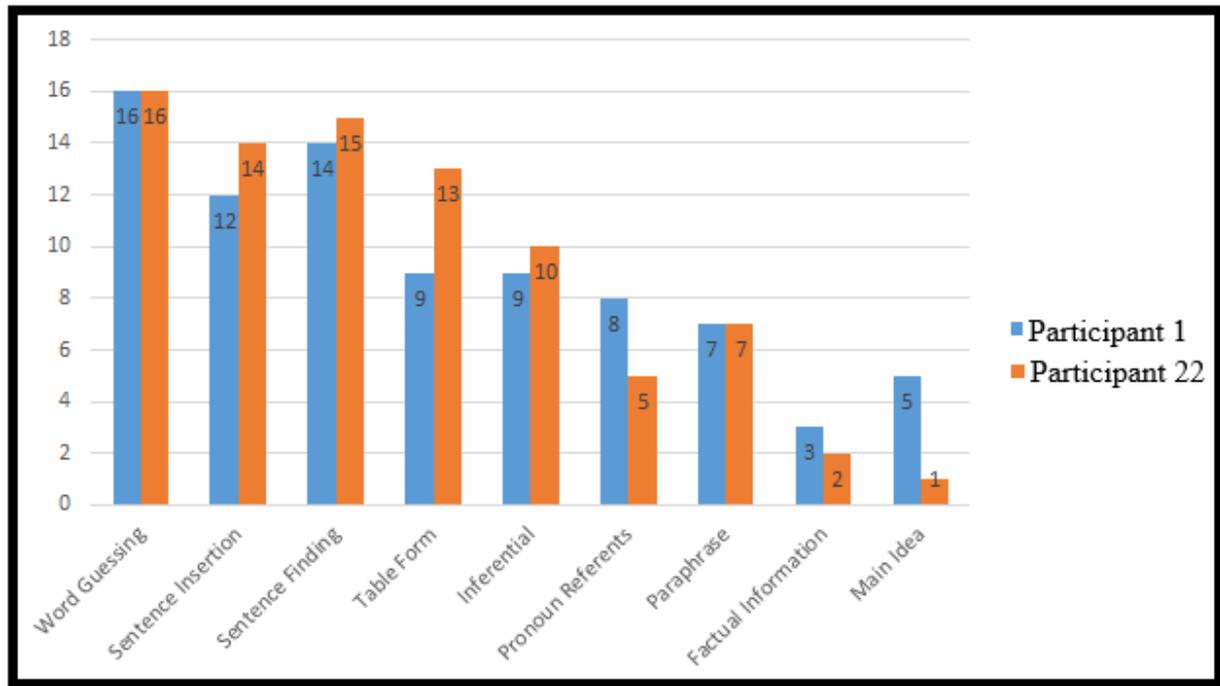
Figure 2. Mediated scores by reading comprehension skills

Figure 2 clearly reveals that even though they performed identically in two skills (word guessing and paraphrase questions), Participant 22's performance was better than his counterpart in the following areas: sentence insertion, where in the passage question (sentence finding), table form, and inferential questions. On the other hand, Participant 1 was stronger in the areas of pronoun referents, factual information, and main idea. This means that the amount of mediation or instruction which should be provided for them varies depending on the specific reading skills; a point which can help teachers with inclusion of different degrees of mediation for different learners in identical items. It seems that mediation required for word guessing and paraphrasing is the same but even a close examination of separate test items might reveal rejection of this idea too (the examination is not included here due to the space constraints).

The performance of Participants 16 and 26 along with Participants 7 and 20 has been also compared with each other, and similar to the previous two participants their levels of required mediation has been examined along with their mediated scores in all nine skills. Participants 16 and 26 who produced the first two lowest scores in the pre-test (10 and 20 respectively out of a maximum of 100) turned out to have an incredibly high learning potential. Figure 3 reveals their improvement under mediation.
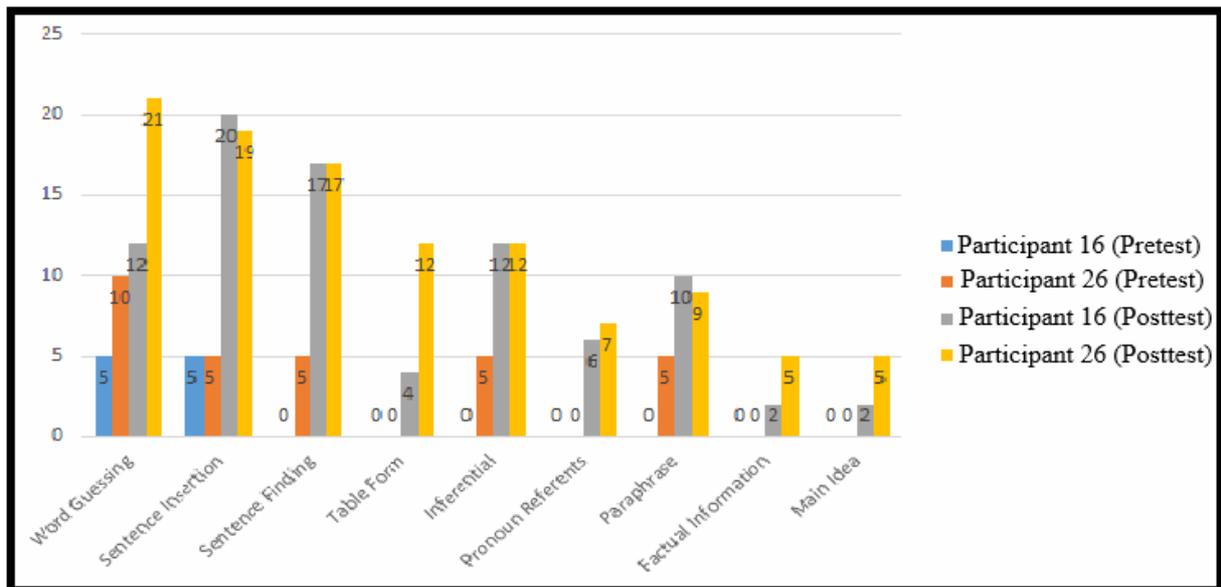
Figure 3. The participants' pre- and post-test scores in each nine reading skills

Note: In case there is no bar, it means the participant received a score of 0 in that skill.

Figure 3 shows both participants' pre- and post-test scores in each nine reading skills. As illustrated, neither of them produced high scores in the pre-test but they unfolded their responsiveness to mediation, which resulted in producing much higher post-test mediated scores as in the participants in Lantolf and Poehner's (2013) study. Grouping the CDRT test items based on the targeted reading comprehension skills showed their more detailed performance. Comparatively, though Participant 16 showed to be equal to Participant 26 in the areas of sentence finding and inferential questions and even better but only in the areas of sentence insertion and paraphrasing, she seemed to be weaker than Participant 26 in the areas of word guessing, table form, pronoun referent, factual information and main idea questions. Instructionally, C-DA is utilized here to uncover and compensate for what traditional testing neglects; based on NDA testing these two participants were not expected to improve but C-DA paved the way for their development. The results revealed that these two participants were actually gainers (to use Budoff's term) because they benefited from the provided intervention markedly (Poehner, 2008; Poehner et al., 2015). Thus, the results were in total discrepancy with the results of Budoff's study, in which some learners were non-gainers in the pre-test and "showed little if any improvement after mediation, performing poorly on both the pre- and post-test administrations" (Budoff, 1987; as cited in Poehner et al., 2015, p. 340).

Therefore, grouping learners simply based on their pre-test scores into high scorers, gainers, and non-gainers would lead to discarding those who can outdo others under graduated and contingent mediation. This would also be in contrast to Vygotsky's opinion

under which understanding an individual's full ZPD exclusively by relying on his/her ZAD is not true. The problem may arise from "lack of fine-grained mediation attuned to the specific needs of individuals", which is one of the "distinct disadvantages of the [interventionist] approach" (Poehner & Lantolf, 2010, p. 318).
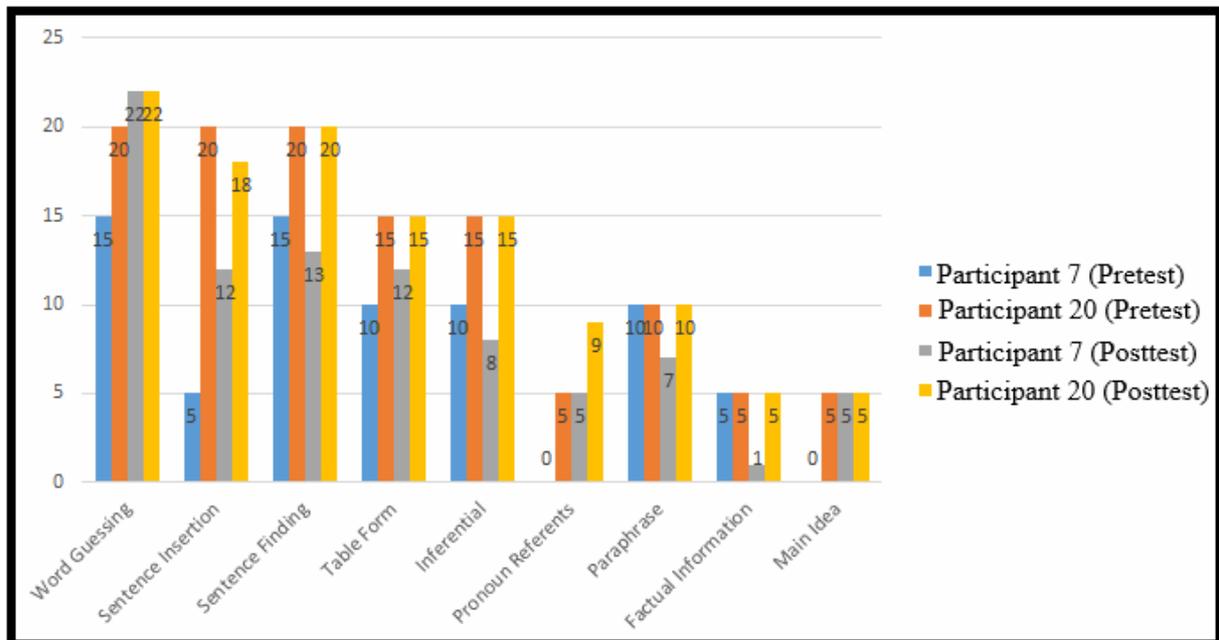

Figure 4. The participants' pre- and post-test scores in each nine reading skills
Note: In case there is no bar, it means the participant received a score of 0 in that skill.

Unlike the ones in the previous Figure, both participants demonstrated in Figure 4 are the two highest scorers of all. Regarding learners of this type, Ableeva (2010, p.120) stated "pre-training [another term for Budoff's high scorers] scores indicate the child's ability to perform on the task independently." The participants' full profile on pre- and post-test performance is illustrated in Figure 4, depicting their high actual (pre-test) scores; especially those of Participant 20. Relatively speaking, Participant 20 who gained only 2 points seemed to have replicated her pre-test scores but that was not the case. She scored higher in the sentence insertion questions on the pre-test (20) but due to her unresponsiveness to mediation she decreased her score to 18 on the post-test. Except for the skills of word guessing and pronoun referent in which she improved her scores (to 22 out of the maximum of 25 and 9 out of the maximum of 10 respectively), her post-test performance was identical to her pre-test one in the other remaining skills.

Regarding Participant 7, it can be observed that, unlike Participant 20 who had identical pre- and post-test scores in some skills, she just improved her scores in some skills,

deteriorated in some others and did not score identically in any skills. For instance, she gained 7, 7, 2, 5, and 5 points for the skills of word guessing, sentence insertion, table form, pronoun referent and main idea questions respectively whereas her scores decreased in the skills of sentence finding, inferential, paraphrase and factual information questions to 2, 2, 3, and 4 points respectively. This backsliding under mediation is also highlighted by Poehner et al. (2015) who explained that even if individuals answered correctly to items, it does not mean that they have not used guessing to reach the correct answer. This is also in line with Vygotsky (1978), who argued that both progressive and also regressive moves are involved in development. Backsliding was experienced by the participants of Ableeva's (2010) study during NDA2 compared to the TA1-IP. In this regard, the software programs for the Transcendence (TR) developed by the researcher shortened reaching the response by guessing through offering explanation for those who score correctly at the first attempt. This is considered as the commitment of DA to supporting learning opportunities by Poehner et al. (2015).

However, readers should be cautioned against getting confused with the gain scores. Earlier it was stated that Participant 7 only gained 13 points under mediation but this Figure shows a lot more than 13; it is because of investigating skills in this nuanced Figure. Items 1, 14, and 15 could be answered by more than one skill and their inclusion increased the number of gained points remarkably. In general, the results are in line with Budoff's proposal under which high scorers had little room for development under mediation owing to their perfect performance on the pre-test. They were also supportive of Poehner et al.'s (2015) study. Optimistically, the future teaching programs should pay attention to high scorers as well since there is no endpoint for development (Poehner, 2008) and producing a high score on a test does not mean lack of flexibility of an individual's level of ability; regardless of whatever it is, as was also stressed by Lidz and Gindis (2003).

In conclusion, counting solely on individuals' pre-test scores to group participants as high scorers, gainers, and non-gainers and subsequently design effective lesson plans might be an insufficient factor (Poehner et al., 2015; and Teo, 2012) as it minimizes the possibility of microgenesis (Lantolf & Poehner, 2013). That is to say, applying DA provides teachers/researchers with a bigger and more nuanced picture of individuals' performance. Hence, in case two learners earn identical scores in the pretest, it does not necessarily imply they have the same proficiency level. Similarly, it is not justifiable to discard low scorers exclusively due to their pre-test performance or consider high scorers as the best performers forever. In this study, the high scorers' trivial improvement under mediation might be due to

applying C-DA, the interventionist approach to DA, which considers "variation across examinees" based on Poehner (2008, p. 25) as "a function of the number rather than the content of the hints, since these are standardized." This means that the shortcoming can be overcome in case both amount and quality of mediation is included over time for individuals.

Practically, this is not possible since the interactionist DA would be beneficial to case study research hence its applicability to large cohorts of individuals is under question (Poehner, 2008). One of the advantages of the interventionist approaches to DA, especially C-DA, is their efficiency, as they provide teachers, researchers, etc. with the opportunity to administer the approach to large numbers of individuals simultaneously and repeatedly (Poehner & Lantolf, 2013; Poehner & Lantolf, 2010). This issue makes application of the interactionist approaches to DA problematic in syllabus design; theoretically, or formal assessment contexts; practically. Besides, owing to its nature of reliance on standardized mediation, C-DA can easily generate each individual's separate scoring profile containing numerical scores which lend themselves easily to psychometric analysis (Poehner & Lantolf, 2010).

5. Discussion and conclusion

This study aimed at identifying the more nuanced impact of scores generated by C-DA on planning a future teaching program; a point which is indicative of the pedagogical implication of the C-DA method. The findings of this study can confirm the practical value of the EFL computerized dynamic assessment procedure through providing in-depth information about various learning needs of the students who have the same standard performance scores. Based on some studies such as Kozulin and Garb (2002) it is confirmed that students with a similar performance level show different, and in some cases drastically different, ability to learn and use new text comprehension strategies. This can confirm the usefulness of DA both in cognitive performance and in such curricular domains as EFL learning.

Through analysis of the obtained results, it was found that C-DA has many advantages. For instance, one of the greatest advantages of the C-DA program is its provision of mediation or intervention when it is required; a point which was underscored by Aljaafreh and Lantolf (1994) who stated that intervention should be provided in gradual progression. In other words, students are provided with hints (mediation) in the C-DA program only if their answers are incorrect. Due to following the interventionist approach to DA in the study and also in order to make the C-DA more systematic, the researchers, in line with Pishghadam, Barabadi, and Kamrorood (2011); Teo (2014); and Shabani (2012), preplanned a series of

mediation (5 hints for each question) which began with the most implicit hints and progressed gradually to the most explicit ones. The C-DA procedure succeeded in overcoming some of the shortcomings of DA approaches such as being time-consuming to administer DA in each class, requiring a fully energetic mediator to take charge of such classes, assessing a smaller number of individual students along with the problem of their age, etc. which there was a general consensus over them by Haywood and Tzuriel (2002), Haywood and Lidz (2007) and Poehner (2008).

Despite such advantages, dynamic assessment has some limitations in its application; that is why it is not used in formal educational contexts a lot. This issue concerned the researchers to take some measures in making DA applicable. For instance, as Pishghadam and Barabadi (2012, p. 73) remarked, "feasibility and concern for psychometric properties of testing are issues that have limited the use of DA approaches." Low number of participants who can be allowed to take part in DA studies and the participants' age are also among its limitations. Haywood and Tzuriel (2002), Haywood and Lidz (2007) and Poehner (2008) all agree upon two more shortcomings of DA: first, it seems it is time-consuming to administer DA in each class and it needs a hyperactive and energetic teacher (mediator) to take charge of such classes. Moreover, DA practitioners worry about its reliability and validity. In addition, since most of the English classes in Iran are large in size, applying the DA procedure, i.e., providing human-to-human mediation to each individual learner, can be unrealistic. Unknown number of instruction sessions or not having "equal time-on-task in DA experimental designs (Nassaji & Swain, 2000, p. 48)" is another problem of DA because in research viewpoint it is preferred to have a certain number showing equal time-on-task. Previously other problems of DA were related to lack of adequate knowledge base and expertise in the field (Haney & Evans, 1999) but due to the increasing interest of some expertise in the field these DA limitations are partially addressed in recent studies.

Furthermore, one of the most important points which still needs exploration is the time which each individual spends on doing a task or test. In the same vein, investigating the relationship between the total amount of time spent on completing a test and the individuals' level of ZPD would either support or reject the argument that those who possess higher ZPD levels require less time to process and perform language activities (Shabani, 2012). In the current study, the overall time each learner spent on responding to all of the items (both in CDRT and in CDRAT) was shown in the scoring file upon completion of the tests, but it was not investigated here because of being far from the aims of the study. Further studies could do so, as well measure the time each learner spends on each item and then examine the

relationship between the overall time and ZPD levels and/or the time spent on each item and ZPD levels.

References

Ableeva, R. (2010). *Dynamic Assessment of Listening Comprehension In Second Language Learning* (Doctoral dissertation). Pennsylvania State University, University Park, USA.

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.

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing Words to Life: Robust Vocabulary Instruction*. New York: The Guilford Press.

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.

Budoff, M. (1987). The validity of learning potential assessment. In C. S. Lidz (ed.), *Dynamic Assessment: An Interactional Approach To Evaluating Learning Potential* (pp. 53–81). New York: The Guilford Press.

Frawley, W., & Lantolf, J. P. (1985). Second language discourse: A Vygotskyan perspective. *Applied Linguistics*, 6, 19-44.

Gibbons, P. (2003). Mediating language learning: Teacher interactions with ESL students in a content-based classroom. *TESOL Quarterly*, 37(2), 247-273.

Haney, M., & Evans, J. (1999). National survey of school psychologists regarding use of dynamic assessment and other nontraditional assessment techniques. *Psychology in the Schools*, 36(4), 295-304.

Haywood, H., & Tzuriel, D. (2002). Applications and challenges in dynamic assessment. *Peabody Journal of Education*, 77(2), 40–63.

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

Kozulin, A., & Gindis, B. (2007). Sociocultural theory and education of children with special needs, from defectology to remedial pedagogy. In H. Daniels, M. Cole & J. V. Wertsch (eds.), *The Cambridge Companion to Vygotsky* (pp. 332-336). Cambridge: Cambridge University Press.

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. (2009). Dynamic assessment: The dialectic integration of instruction and assessment. *Language Teaching Journal*, 42(3), 355-368.

Lantolf, J. P. (2000). *Sociocultural Theory and Second Language Learning*. Oxford: Oxford University Press.

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.

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., & Thorne, S. L. (2006). *Sociocultural Theory and the Genesis of Second Language Development*. Oxford: Oxford University Press.

Lidz, C. S., & Gindis, B. (2003). Dynamic assessment of the evolving cognitive functions in children. In A. Kozulin, B. Gindis, V. S. Ageyev, & S. M. Miller (eds.), *Vygotsky's Educational Theory in Cultural Context* (pp. 99-116). Cambridge: Cambridge University Press.

Mardani, M., & Tavakoli, M. (2011). Beyond reading comprehension: The effect of adding a dynamic assessment component on EFL reading comprehension. *Journal of Language Teaching and Research*, 2(3), 688-696.

Merriam, S. B. (1997). *Qualitative Research and Case Study Applications in Education*. San Francisco: Jossey-Bass Publishers.

Nassaji, H., & Swain, M. (2000). A Vygotskian perspective on corrective feedback in L2: The effect of random versus negotiated help on the learning of English articles. *Language Awareness*, 9(1), 34-51.

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. (2008). *Dynamic Assessment: A Vygotskian Approach to Understanding and Promoting L2 development*. Pennsylvania State University, University Park, PA: Springer Science & Business Media.

Poehner, M. E. (2005). *Dynamic Assessment of Oral Proficiency among Advanced L2 Learners of French* (Unpublished dissertation). Pennsylvania State University, University Park, PA, USA.

Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, 9, 233-265.

Poehner, M. E., Zhang, J., & Lu, X. (2015). Computerized dynamic assessment (C-DA): Diagnosing L2 development according to learner responsiveness to mediation. *Language Testing*, 32(3), 337-357.

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.

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Adolescent Literacy*, 78(1), 40-59.

Smagorinsky, P. (1995). The social construction of data: Methodological problems of investigating learning in the zone of proximal development. *Review of Education Research*, 65(3), 191-212.

Snow, C. E. (2002). *Reading for Understanding: Toward an R & D Program in Reading Comprehension*. Santa Monica, CA: Rand.

Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic Testing*. New York: Cambridge University Press.

Summers, R. (2008). *Dynamic Assessment: Towards a Model of Dialogic Engagement* (Doctoral dissertation). University of South Florida, Tampa, Florida, USA.

Tajeddin, Z., & Tayebipour, F. (2012). The effect of Dynamic Assessment on EFL learners' acquisition of request and apology. *The Journal of Teaching Language Skills*, 4(2), 87-118.

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*. Cambridge, MA: Harvard University.

Wells, G. (1998). Using L1 to master L2: A response to Anton and DiCamilla's 'socio-cognitive functions of L1 collaborative interaction in the L2 classroom'. *Canadian Modern Language Review*, 54(3), 343-353.

Wertsch, J.V. (1985). *Vygotsky and the Social Formation of Mind*. Cambridge, MA: Harvard University Press.