Research Article

Analyzing the Impact and Opportunity of Computer-Based Dynamic Assessment on the Development of Understanding Prepositions among Iranian EFL Learners

Alireza Mohammad Beigi [1], Reza Rezvani [2], Saeed Yazdani [3]

1, 3. Assistant Professor, Department of English, Bushehr Branch, Islamic Azad University, Bushehr, Iran
2. Associate Professor, Department of English, Shiraz University, Shiraz, Iran

* Corresponding author: **Reza Rezvani,** Email: rezvanireza8620@gmail.com

ARTICLE INFO

ABSTRACT

The current study examines the effect of computerized dynamic assessment (C-DA) on Iranian EFL learners' prepositions. More specifically the aim of the study was to investigate the difference between dynamic and non-dynamic tests and to investigate how mediation works in terms of the learning potential scores of high and low achievers with the help of C-DA. To this end, 30 learners of Iran Language Institute (ILI) at Shiraz branch were selected from two intact elementary classes. They were randomly given two test formats. On the whole, the two groups were given the tests in the computerized format, one in the dynamic format and the other one in the non-dynamic format. To achieve the objectives, computer software was developed. The computerized test included 30 items and five hints presented for each item in case the learners could not answer the item correctly. Analysis of findings showed that the computerized dynamic test had a positively important impact on the development of students' knowledge of prepositions and getting some information about their learning potential. It was further determined that the utilization of computerized dynamic assessment may simultaneously lead to the enhancement of learners' ability, thereby providing educators with a clear picture of learners' potential for learning. Given the significance of C-DA in revealing learners' learning potential score (LPS), The findings of the study have significant implications for educators, as they can use these findings to design distinct individualized learning programs that address the learning requirements and capabilities of their students.

Introduction

The origin of dynamic assessment is founded in the theoretical conceptualizations of Vygotsky (1978). According to Chaiklin (2003) this conceptualization has been interpreted differently by different scholars in different contexts. Vygotsky (1987) provided a definition for the zone of proximal development, which refers to the gap between the current level of development achieved through independent problem-solving and the level of potential development that can be reached through problem-solving with the guidance of adults or in collaboration with more influential individuals. The zone of proximal development (ZPD) assumes a critical role in the advancement and implementation of online computerized interventionist DA procedures. Among DA scholars, Lantolf and Poehner conducted a huge project for assessing reading and listening skills through computerized dynamic assessment for five languages of Russian, Chinese, French, and German. In recent years, quite interestingly, many Iranian TEFL researchers have become interested in doing DA studies. A big portion of these studies are devoted to conducting to the computerized dynamic test of different aspects of English. Barabadi (2010) and Mehri Kamrood (2011) were among the first researchers who conducted C-DA test for assessing the reading and grammatical ability of English language learners.

Ebadi and Saeedian (2016); Modarresi and Alavi (2014); Nasiri, Vahid Dastjerdi and Tavakkoli (2013, among others), have also developed special software in order to assess reading, grammar, and transition words or phrases. The general problem that this study is going to address is the inefficacy of traditional (nondynamic) tests of grammar in depicting a completely representative picture of the learners' grammatical abilities. This problem will be solved by developing a computerized dynamic test of grammar (CDTG). Moreover, an inherent problem of nearly all L2 DA studies done so far is the narrowness of their perspective regarding the number of individuals and subjects and the aspects of the language that are being tested (e.g., Ableeva, 2008; Anton, 2009; Birjandi &Ebadi, 2009; Lantolf & Poehner, 2008; Kozulin & Garb, 2001). This study is an attempt to address this problem by developing a computerized (interventionist) DA that will have the potential to be administered to as many students as possible with a wide range of grammatical aspects involved.

Another reason for conducting this research is that most University EFL learners, no matter how competent they are in other areas of language proficiency such as reading, writing, speaking, etc., have problems dealing with grammar tests, in terms of answering questions such as error identification type questions in which their metalinguistic knowledge of English is involved. Though not used in recent TOFEL IBT and IELTS, these kinds of questions are still used in most language proficiency tests both locally and internationally, e.g., almost all language proficiency tests developed for entering the PhD programs in Iranian universities; First Certificate in English Language (FCE) in Cambridge University, the Comprehensive English Language Test (CELT), and the Community English Program Placement Test in Teachers College, Columbia University. Through providing graduated and contingent feedback, this study aims to help the learners exert their full potential (both the ZAD and ZPD) in answering the questions while they develop their grammatical knowledge at the same time.

The main goal of this study was to investigate how a computerized dynamic assessment of prepositions could potentially enhance learners' comprehension of L2 prepositions. Additionally, another crucial aspect of this study was to explore the ability of DA to assess learners' ability to learn. Consequently, the present study aims to address the following research questions:

1. Is there any meaningful difference in the scores obtained by learners in computerized dynamic assessment and computerized non-dynamic assessment?
2. Is computerized dynamic assessment effective in uncovering examinees' potential for learning?
3. Does the learning potential score discriminate between learners who have the same actual

score on non-dynamic English grammar knowledge test?

4. How does the zone of potential development (ZPD) of high and low achievers differ through computerized administration of hints?

Limitations of the Study

The main and the most obvious limitation of this study which is a drawback for any DA study that follows an interventionist approach is that it cannot tailor the mediation to the personal needs of each individual learner. Indeed, they have access to prefabricated conventional hints and feedback that may not tap into particular problems that students face at the time of taking the exam, even though the hints are organized from the most implicit to the most explicit. Accordingly, there is always the possibility of designing other mediation schemes that are more in line with the original ideas of Vygotsky. A major limitation of the present C-DA study like those of Poehner and Lantolf (2013) and Poehner et al. (2015) is related to the format of the test items. So far, the only test format that has lent itself to prerequisites of the C-DA projects is the multiple-choice format. Other formats are very difficult to be called upon in C-DA projects regarding what we have at our disposal technologically speaking. Moreover, another issue is related to the number of choices for each item which is usually five choices. This format may not follow the standard format of the traditional tests due to the interactive nature of the test item. In other words, items should have the capacity to be tried as many times as needed in the dynamic test and not reveal the correct choice. Thus, the researcher needs to manipulate the so-called standard test items and raise the number of options to five in order to meet the requirements of a C-DA project. Another important limitation of this study is related to the cost and time of developing such online C-DA studies that are directly related to the quality of the software and the website that is going to be developed and established. High-quality C-DA software needs high-quality experts in programming to be hired which in turn would be very costly. Studies like these need to be developed, piloted and commented on many times so that they have the least loopholes at their final stage of conducting. Finally, the sampling procedure would be convenient sampling which will reduce the generalizability power of the findings to a larger population. The researcher will not follow the interactionist version of dynamic assessment due to the nature of all C-DA procedures that require pre-planned and prefabricated hits and prompts. Moreover, this study will lose a large number of its to-be participants due to the lack of technological means such as PCs and access to the internet, etc.

Dynamic Assessment (DA):

"Dynamic assessment is an approach to understand the individual differences and their implications for instruction that embeds intervention within assessment procedure" Lidz and Gindis (2003, p. 99)." However, Hasson and Joffe (2007) define DA as "a range of methods and materials to assess individuals' potential for learning. It aimed to reveal the maximum level of performance by providing mediation in the course of assessment session" (p. 10).

Computerized Dynamic Assessment (C-DA):

A new kind of DA in which a computer takes charge of the mediation process. "C-DA has several distinct advantages, including the following: it can be simultaneously administered to large numbers of learners; individuals may be re-assessed as frequently as needed; and reports of learners' performances are automatically generated" (Poehner, 2008a, p.177).

Non-dynamic/Static Assessment (NDA)

"Static tests grounded in psychometric principles, assume that a person's solo performance on a test represents a complete picture of the individual's capabilities" (Poehner & Lantolf, 2005, p.234).

Zone of Proximal Development (ZPD):

Zone of proximal development is defined as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem-solving under adult guidance or in collaboration with more capable peers" (Vygotsky, as cited in Aljaafreh & Lantolf,1994).

Interventionist DA:

In the interventionist approach to dynamic assessment, the mediation takes place with the mediator using prefabricated hints, suggestions, and explanations in a preplanned order to help the learners while they move through each part of the test.

DA and the Psychometrics

Among DA researchers, there is always a debate about how they should react to the psychometric issues which are still the mainstream standards in modern testing. Some of them such as Haywood and Lidz (2007); Tzuriel (2003); Grigorenko and Sternberg (2001) believe that DA strategies can be employed as complementary to NDA tests in order to provide test users with a complete picture of test takers' abilities. Some other more radical DA researchers such as Lantolf and Poehner (2008) believe that viewing DA as a supplement to NDA testing is against Vygotsky's original ideas about unifying assessment and instruction into a seamless activity. In other words, they believe that DA procedures can replace NDA testing. In this section, some psychometric considerations such as reliability, validity, and generalizability are discussed in relation to DA.

Reliability of DA

According to Lantolf (2009), psychometrically speaking, reliability follows the ontology of human being as an autonomous individual. In other words, the environment does not play an important role in human beings' development; hence, while assessing learners, the outside factors should be controlled in order not to contaminate the results of tests. He believes that Vygotsky's theory of human being development is in sharp contrast to that of the psychometricians. Poehner and Lantolf (2007) believe that effective assessment is characterized by change rather than stability (Lantolf, 2009). Furthermore, Lantolf and Poehner (2004) point out that standardization is at the heart of reliability, so it seems that DA procedures that follow an interventionist approach are more in line with reliable static tests. However, there exists a much more important issue in all DA procedures which runs counter to the very first feature of any reliable test, that is, the neutrality of the assessor. Lantolf

and Poehner (2008) point out that while the assessors' intervention during a dynamic test is indispensable; in an NDA test, it is a powerful source of test method effect which contaminates the reliability of DA tests. Consequently, no DA test is reliable from a psychometric perspective.

Validity of DA

DA researchers have their own stance in that while reliability is by no means achievable in DA tests, validity can be brought up and discussed, of course, through somehow different orientations towards the notion of validity (Lantolf & Poehner, 2004). As for construct validity, Lantolf (2008a) claims that DA researchers must define their constructs and also argue for the validity of employing their procedures and instruments. However, there is another important issue that is very critical in conducting any DA procedure. Lantolf and Poehner (2008) state that, irrespective of the specific construct, any DA procedure should be strongly concerned with the general construct of "development". They believe a DA procedure is truly effective when it promotes development during its administration. That is to say, a DA test has construct validity to the extent it promotes development. In other words, Poehner (2008a) claims that DA tests are development-referenced because their effectiveness is contingent upon their influence on learners' development.

Literature Review

Orikasa (2010) conducted a case study of interactionist dynamic assessment in second language learning context by tutoring English oral communication. The aim of this study was to investigate how interactions between a mediator and an L1 Japanese student are negotiated and help develop the learner's performance. The findings of the study showed that interactionist dynamic assessment in the second language context is effective in helping the learner overcome problems and perform better through negotiated interactions with the mediator and revealing the learner's actual competence. In a different study with a different scope, Alirezaei and Ghanbarpour (2016) shifted their attention from the examinee to the examiner. This study used methodical procedures for qualitative meta-synthesis of the target research

domain to synthesize available quantitative and qualitative primary research reports. The discursive reading and systematic review of research findings and discussions across study reports revealed a primary theme, that is, 'dynamic assessment with a shift in focus from post-positivism to pragmatism, and a couple of secondary themes 1) classroom-based second language dynamic assessment as a post-achievement test condition and 2) mediators' sense of accountability that can provide a refined worldview on dynamic assessment. It also revealed the relevance of assessment to classroom context and an informed understanding of the ascendency of mediators over the effectiveness of dynamic assessment. Examining the criticism of dynamic assessment, Hosseini and Ghonsooly (2017) adopted a design in which instruction and assessment were integrated. After presenting a battle of views on DA, they found out that the concept has two different faces, with one aspect being attractive, encompassing ethical principles, justice, and societal equality, and the other aspect of DA being characterized by adaptability, interventionism, and stability. The researchers then concluded that each system should be criticized by specific criteria. Similarly, in a study on developing EFL learners' speaking skills through dynamic assessment, Ebadi and Asakereh (2017) studied beginners and advanced learners. Building on the theoretical underpinnings of dynamic assessment. The improvement of speaking abilities was the focus of investigation for these researchers, who explored the impact of dynamic assessment. For the investigation, an inexperienced and a skilled learner of the English language were selected as the participants to accomplish this goal. The participants were given instructions to create a series of visual narratives, with guidance tailored to their Zone of Proximal Development (ZPD), in order to collect the necessary data. The analysis of the data was conducted using micro-genetic and thematic analysis as a framework, to determine any potential changes in the participants' cognitive development The researchers regarded the participants' private speech as a sign of their shift from external influence to personal regulation. The findings suggest a significant improvement in the participants' cognitive skills. Moreover, the results of the thematic analysis carried out on the

unstructured interviews revealed that the participants expressed satisfaction with the intervention. DA has been welcomed warmly in the Iranian context. Researchers believe that the novelty and the attractiveness of DA have attracted many Iranian TEFL students. Hashemi, Ketabi, and Barati (2015) conducted a group-DA study that investigated the listening comprehension of learners in three proficiency levels. They followed the sandwich format of DA, that is, they conducted a pretest-mediation (enrichment)-post-test procedure. Their findings revealed that G-DA is able to both determine and support the development of listening comprehension of learners. Rajaeizadeh, Biria, and Kheirzadeh, (2015) examined the instructional effectiveness of DA on English vocabulary learning of young Iranian EFL learners. They followed the interventionist cake model of DA. The researcher provided the learners with hints and prompts when they were taking the test. The mediators used an observation chart in order to record the behavior of learners during the first test. Then, they administered a parallel test two days later to see if there was a difference between learners' performance on the first and the second test. In order to check for the lexical recall of the language learners, they conducted two near and far transcendence tasks. The results of their study revealed that learners' English vocabulary ability was promoted.

Also, Taheri (2018) conducted research, seeking to reconcile the assessment and instruction dualism by integrating them into a single unified entity. Twenty-eight TEFL students participated in this study. Microgenetic analysis and idea unit analysis were used to reveal the frequency and quality of mediational moves across the assessment sessions. The findings of the study indicate that dynamic assessment allows establishing the actual level of learners' listening ability diagnosing the potential level of their listening comprehension, while at the same time promoting their comprehension. Zangoei et al. (2019) conducted a study to examine the impacts of an interventionist Computerized Dynamic Assessment (CDA) on L2 pragmatic comprehension, specifically focusing on speech acts, routines, and implicatures. The researchers discovered that the implementation of

computerized DA instruction had a noteworthy positive effect on the learners' pragmatic comprehension. In a separate study, Rezai et al. (2022) investigated the consequences of online peer-DA on the writing skills of Iranian EFL learners. They observed a significant improvement in the participants' writing skills following the interventions.

Knowledge of prepositions

Language acquisition is a fascinating aspect of human development, and the role of grammar is vital in the construction of sentences in a specific language. When it comes to English grammar, the correct utilization of prepositions holds significant importance as it ensures the appropriateness, comprehensibility, and coherence of a sentence. Prepositions in the English language are commonly defined as words that establish connections between nouns and other words, thereby indicating their relationship. My research will center on the premise that the acquisition of prepositions is enhanced through direct instruction for L2 learners, given that English is the language they are acquiring.

Learning a language is about enhancing skills and working knowledge of its linguistic features, including phonology, lexis, syntax, semantics, and graphology. Those features involve the language components that need to be acquired by language learners to be proficient language users. However, understanding them is not an easy task. It still becomes a challenge for many English foreign language (EFL) learners especially when they are dealing with English prepositions. A preposition demonstrates a relationship from one word to another in the sentence. According to Walker (1982, p.123), the preposition demonstrates how a noun or pronoun, and sentence parts are connected.

Learning Potential Score (LPS)

Among the scholars of DA Kozulin & Garb (2002) conducted a small-scale study of dynamic assessment of EFL adult text comprehension. They assessed the students' ability to learn and use reading comprehension strategies following the test-teach-retest (sandwich) DA format. Students took a static test. Then the test was previewed by both

teacher and students together focusing on the strategies that were effective for each item. A post-test served to assess whether students had benefitted from the mediation.

The findings of their research validated their assertion that a DA approach examines not only the prospective knowledge of learners, which may be significantly influenced by cultural differences but also emphasizes the learners' ability to take advantage of mediation. The findings revealed that there was an enhancement in the abilities of many learners in the post-test examination. But not all. Kozulin and Garb devised a method for the

$$LPS = \frac{(S_{post} - S_{pre})}{S_{max}} + \frac{S_{post}}{S_{max}} = \frac{2S_{post} - S_{pre}}{S_{max}}$$

implementation of the measurement of students' potential for learning that distinguishes between students with high and low learning potential. They emphasized that certain students with both high and low learning potential attained equivalent scores in the initial examination, accordingly demonstrating that dynamic assessment adds significant information that is hidden during static testing. In order to operationalize the students' learning potential, they developed the following scoring method:

According to Kozulin and Garb (2002, p.121): "the learning potential score (LPS) has to represent both gains made by the learner from pre- to post-test and achievement score at the post-test. Where S pre and S post are pre-and post-test scores respectively, and S max is a maximal obtainable score. The above formula provides a theoretical basis for distinguishing between high learning potential and low learning potential students." In their study, Kozulin and Garb (2002) claimed that a student who had a low score = 50 on a pre-test, but significantly progressed and reached the maximum score of 100 on the post-test would get a very high LPS of 1.5; on the other hand, a student who got a low score of 50 at a pre-test and made no progress after the mediation session getting the same score of 50 in the post, would end with a very low LPS of 0.5. All other students' scores can be interpreted in the same manner.

Later they concluded that their learning potential score can add important information about the student's learning ability which goes unnoticed when a standard non-dynamic test is

administered. They also claimed that the results of their study confirm the fact that DA cannot be limited to the realm of cognitive performance. Other curricular areas such as EFL learning could lend themselves to DA procedures. They also emphasized the fact that the quality of mediation can be very critical in that a different meditational style can result in a different pattern of LPS for the same learners.

They also stated that LPS score can provide us with valuable information that can be used for developing learning plans for individuals with different learning needs. They further claimed that for example for a learner with an average pretest score and a low learning potential score the assessor should prepare information processing strategies, that is he should teach them how to learn. More challenging materials should be provided for a learner with an average pretest score and a high LPS. In the case of a low pretest score and a low LPS, the learner should be provided with intensive courses in general learning and problem-solving skills based on simple EFL materials.

Method
Participants
The subjects of the study included 30 Iranian learners of Iran Language Institute at Shiraz branch. They were members of two independent classes. Also, they were all elementary learners of English attending ILI to learn English language. As such, they were randomly given two tests. In the experimental group, the participants underwent a computerized dynamic test while the control group received the non-dynamic test of preposition. On the whole, the two groups were given the tests in the computerized format, one in dynamic format and the other one in the non-dynamic format.

Instruments
As this study utilizes an experimental format, we employ software that possesses the ability to assess learners' knowledge of prepositions by providing predetermined hints in the event of a mistake. The preposition questions were derived from the error detection segment of the publication titled "Test your prepositions" authored by Jones and Allsop in 2000. In this manner, every class was presented with a set of 30 homogeneous preposition questions.

Test Piloting
The main goal of conducting test piloting was to gather feedback from both professionals and students about the format and information covered in the C-DA test. Poehner et al. (2015) carried out the test piloting process under the supervision of the researcher. That is, a total of 50 test items were initially chosen, and hypothetical hints on the basis of the Zone of Proximal Development (ZPD) were organized for each individual test item, following with the regulatory scale for presenting ZPD-based feedback proposed by Aljaafreh and Lantolf (1994). In the next stage, he asked experts in dynamic assessment regarding their evaluations of the test items and hints. Following this, the researcher proceeded to administer the test to 25 learners without any external intervention in order to get information regarding the proficiency level and response tendencies of learners, along with the difficulty of individual items and the assessment as a whole. After analyzing learners' responses, the total number of items was decreased to 30, resulting in the exclusion of 20 items due to various reasons. One obvious reason related to the simplicity of items. In other words, some of the items were too simple for the majority of learners, thus they were set aside as they hindered the initiation of the mediation process. Finally, the ultimate version of the examination, consisting of the items along with the hints, underwent a review process conducted by two professors from Shiraz University, resulting in some modifications being made.

Test design and construction
Grammar items were chosen from the book *Test your preposition* authored by Jones and Allsop (2000). The most advantageous structure for our test was the error identification type items, as this particular structure had the potential to be accompanied by helpful hints for each question. Multiple choice items were deemed inappropriate in this context due to the increased likelihood of guessing. Additionally, with each incorrect answer, the number of choices would decrease, further increasing the probability of guessing the correct answer even without utilizing the provided hints.

After providing the aforementioned items, a total of five hints were presented. In line with the Regulatory Scale (1994) employed by Aljaafreh and

Lantolf, the hints were ascertained to progress from implicit to explicit in nature.

Three of the hints were organized to assist the learners in identifying the error part. Since the learners were unable to detect the mistake after receiving the third hint, the fourth hint would present them with the erroneous section, while also requiring them to write the accurate version of the incorrect part. If they provided the correct response, they would move on to the next item; otherwise, they would receive the final hint. Generally, the purpose of the initial hint was simply to inform the test takers that their answer was incorrect, thus enabling them to refer back to the item and make another attempt. The second hint describes the nature of the error, while the third hint is basically the scope of the error that was highlighted. In certain instances, the sequence of presenting the hint was modified. These alterations in order were motivated by the intention to present the hints in an increasingly explicit manner, moving from implicit to explicit.

In the fourth hint, the focus was solely on the incorrect part, requiring learners to produce the correct version of the incorrect part. In the last hint, learners were presented with the correct response along with a concise explanation of the grammatical issue in question. This explanation was subsequently accompanied by an example that demonstrated the appropriate application of the specific point. This emphasizes the idea that the main objective of DA studies is not only to evaluate the learner's understanding but also to enhance their comprehension of the aforementioned aspect.

Data collection procedure

In the initial stage, the computerized test comprised 30 items, with each item being accompanied by five items in the case that the learner was unable to respond to the items correctly. Subsequently, After the test has been finished, a file containing the scores is produced with specific and detailed information. For every learner, two scores are provided, namely dynamic and non-dynamic. Additionally, the scoring file includes information on the number of hints utilized for each question, as well as the whole time spent on the examination. For better understanding, the general scheme of presenting hints is presented below:

Q: In ten o'clock in the morning, I went to the beach to catch fish with my net in Sunday.
Hint 1: Your answer is wrong! Try again.
Hint 2: There is a prepositional error.
Hint 3: The error is in the highlighted part.
Q. In ten o'clock in the morning, I went to the beach to catch fish with my net in Sunday.
Hint 4: The highlighted part is a prepositional error. Try to rewrite the correct form.
Hint 5: The preposition should be changed proportionate to time, place, etc.
Eg; in sunday to → on Sunday
Figure 1 displays screenshots of the initial English grammar test along with the presented mediations. Additionally, Figure 2 illustrates the performance of learners on the English grammar test, which was evaluated using a scoring profile that not only reported the score achieved with the use of hints but also the score achieved without any hints.
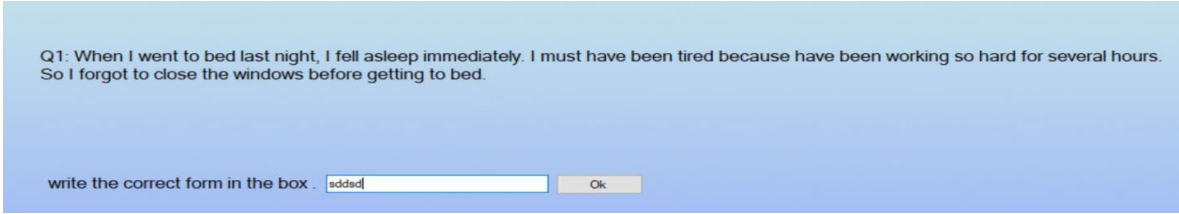
Q1: When I went to bed last night, I fell asleep immediately. I must have been tired because have been working so hard for several hours. So I forgot to close the windows before getting to bed.

write the correct form in the box . [sddsd]   [Ok]

Figure 1. *Screenshots of a question and its ZPD-based mediation*

"Name: ....Mohsen.... .... "　　　　"Age: 20
" Institute: Iran Language Institute "
"Gender: Male"
"Score gained with the use of hints: 35"
"The total number of hints used: 84"
"Score gained with no hint: 10"
"Total time spent::41:48"
"The number of hints used 1: 1 "
"The number of hints used 2: 4 "
"The number of hints used 3: 2 Missed"
"The number of hints used 4: 0 "
"The number of hints used 5: 0 "
"The number of hints used 6: 3 Missed"
"The number of hints used 7: 0 "
"The number of hints used 8: 0 "
"The number of hints used 9: 0 "
"The number of hints used 10: 0 "
"The number of hints used 11: 1 "
"The number of hints used 12: 3 "
"The number of hints used 13: 0 "
"The number of hints used 14: 0 "
"The number of hints used 15: 2 "
"The number of hints used 16: 0 "
"The number of hints used 17: 0 "
"The number of hints used 18: 0 "
"The number of hints used 19: 0 "
"The number of hints used 20: 0 "

"The number of hints used 21: 0 "
"The number of hints used 22: 0 "
"The number of hints used 23: 0 "
"The number of hints used 24: 4 "
"The number of hints used 25: 0 "
"The number of hints used 26: 0 "
"The number of hints used 27:4 "
"The number of hints used 28: 3 "
"The number of hints used 29: 3 "
"The number of hints used 30: 2 "

Figure 2: A learner's scoring profile

Data analysis

To ascertain the statistical significance of the differences between the mean scores of the experimental (dynamic) and control (non-dynamic) groups, we employ the independent sample t-test. Moreover, the eta squared statistic is utilized (Dornyei, 2007) to figure out the strength of this difference. Accordingly, we employed the Kosulin and Garb (2002) formula to examine the potential score of the learners (LPS).

The test taker's personal information

1. Non-dynamic score

The score is determined based on the initial attempt made by the learners for each question, without considering the number of hints they may have received. Nevertheless, to compare this score with the test's dynamic score, it is measured on a scale from 0 to 100 points, with each question worth five points. For instance, a learner (referred to as Ali, but using a pseudonym) who correctly answered five questions without using any hints received a non-dynamic score of 25.

2. Dynamic score

The true dynamic score of individuals taking the test is reflected by the number of hints they use. Considering the fact that there are 150 hints accessible for each examination, with five hints allocated to each item, it is possible to determine their dynamic score by deducting the number of hints they acquire from the total number of hints.

Results

This investigation represents an endeavor to construct and execute an interventionist C-DA of knowledge of prepositions for Iranian EFL learners. The quantitative aspect of this investigation involved the implementation of various statistical procedures on the collected data. Before the main examination, an initial evaluation was conducted to determine the normality of the participants' distribution. As illustrated in Table 1 below, the Sig. The Kolmogorov-Smirnov value of .061 suggests normality in terms of the pretest. In simpler terms, a non-significant result suggests a normal distribution. In cases where the assumption of normality holds, ANOVA, a parametric statistic, was employed for the current study.

Table 1.
Test of Normality

	Kolmogorov-Smirnov^a			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	Df	Sig.
Pretest	.128	45	.061	.908	45	.052
a. Lilliefors Significance Correction						

With respect to the first research question, Table 2 shows the descriptive statistics of the dynamic (experimental group) and non-dynamic (control group) scores of the participants.

Table 2.
Descriptive Statistics of Participants' (N=30) Scores

Groups	Mean	Std. Deviation	Variance
Experimental	96.07	20.99	440.61
Control	33.12	17.98	323.41

As can be seen from Table 2, providing the learners with graduated hints through a dynamic test has contributed more to the improvement of the learners' scores. While the mean performance score of the participants on the non-dynamic test is almost 33, that of the learners' dynamic score has increased to 96. This shows that the test takers have outscored in the dynamic test. However, the participants' non-dynamic scores are more homogeneous than those of DA. In order to calculate the statistical significance of the difference between these sets of scores, an independent sample t-test was performed. The results as appears in Table 3, illustrate that there was a meaningful difference between the DA and NDA scores ($t(63) = 34.3$, P. <.001).

Another way of determining whether participants in the study actually gained advantages from mediation can be achieved through the computation of the effect size. There are different

formulas for calculating the effect size; Dorniey (2007) suggests the "eta squared" formula for calculating the effect size. The effect size is low if it varies around .01. It is moderate if it varies around .06 and the effect size is high if it varies around .14. Regarding the "eta squared" formula, our sample effect size was .9 which was considered a large effect.

Table 3.
Paired Sample T-Test for DA and NDA Scores

Type of the test	Mean (SD)	df	t	P
Dynamic	96 (20.99)	63	34.3	.001
Non-dynamic	33 (17.98)			

Another way of computing the effect size is through Pearson r correlation. In this particular situation, the range of the effect size can span from -1 to +1. Cohen (1992) asserts that the effect size is deemed as low when the value of r fluctuates around 0.1. If r is approximately 0.3, the effect size is categorized as medium, whereas if r deviates by more than 0.5, it is referred to as large. The correlation between dynamic and non-dynamic scores in this specific investigation is (.72), indicating a large effect size. The second question of the study specifically examined the capacity of Dynamic Assessment (DA) to assess the extent of learners' Zone of Proximal Development (ZPD). We provide information on the potential for students to enhance their learning and progress in the future. This distinguishes DA from traditional tests. We used Kozulin and Garb's formula to see how well DA helps students learn and improve in the future, considering how well they can learn with help. Let us now consider how the Learning Potential Score (LPS) of the learner is computed:

Ali's NDA score: 20

His DA score: 50

The highest DA score on the exam: 93

$$LPS = \frac{(S_{post} - S_{pre})}{MaxS} + \frac{S_{post}}{MaxS}$$

In this study, S pre and S post represented the non-dynamic and dynamic scores, respectively. Additionally, Max represented the maximum attainable score which amounted to 93 in the provided instance.

Ali's LPS $= \frac{(50-20)}{93} + \frac{50}{93} = .86$

As highlighted previously, this study was concerned with the potential ability of DA measures, compared to non-dynamic measures, to lead to a better understanding of the students' individual learning ability. In other words, it also Table 4.

addressed the capability of dynamic tests to assess the size of learners' ZPD, and described the learners' ever-changing and differing ability with assistance. LPS is found specifically effective when it aims to differentiate among learners with identical NDA scores. To achieve this, Kozulin and Garb (2002) formula in terms of LPS was calculated.

In Table 4, the learners' LPSs exhibited a range of values from .27 to 1.26. LPS emphasizes the fact that the progress made by the participants in their performance on the dynamic test was not the same. Consequently, by means of this score, it became feasible to distinguish between test takers who possessed an identical NDA score. Learners who exhibited significant progress from the non-dynamic to the dynamic test exhibited a high LPS, whereas those who made slow progress displayed a low LPS. Once again, consider the test taker mentioned previously, who obtained an LPS of .86 on the exam. Henceforth, through this score, it became feasible to differentiate between examinees that possessed identical NDA scores. Those individuals who demonstrated noteworthy advancement from a non-dynamic to a dynamic examination exhibited High LPS values, whereas those who exhibited low progress achieved lower LPS values. Once more, let us consider the aforementioned examinee with an LPS of .86 on the exam. Another learner, who possessed an identical static score of 25, exhibited a learning potential of .67. Consequently, the two learners exhibited differently in terms of potential for learning despite having an identical non-dynamic score. Stated differently, two learners who obtained a non-dynamic score of 25 displayed different levels of advancement on the dynamic assessment. One examinee attained an LPS of .86, whereas the other achieved an LPS of .67.

Descriptive Statistics of Test Takers' LPS on the Test

	N	Minimum	Maximum	M	SD
The test	30	.27	1.26	1.21	1.3

The participants' LPS in the study varied between 0.27 and 1.26. To answer the third research question which asked whether LPS possesses the ability to differentiate between learners who have identical NDA Scores, we conducted a comparative analysis of ten learners who all obtained an NDA score of 30. Figure 1 provides a clear description of the considerable distinctions among these ten students regarding their potential learning scores.

For example, Students number 2 and 4 in our list had the same NDA of 30 but their LPS scores were 0.97 and 1.18 respectively. This indicates that while from the standpoint of a non-dynamic test, the prepositional knowledge of all these students is equal, the learners' LPS and subsequently their dynamic score could be distinguished among them by assessing the ZPD as well as their zone of actual development (ZAD).
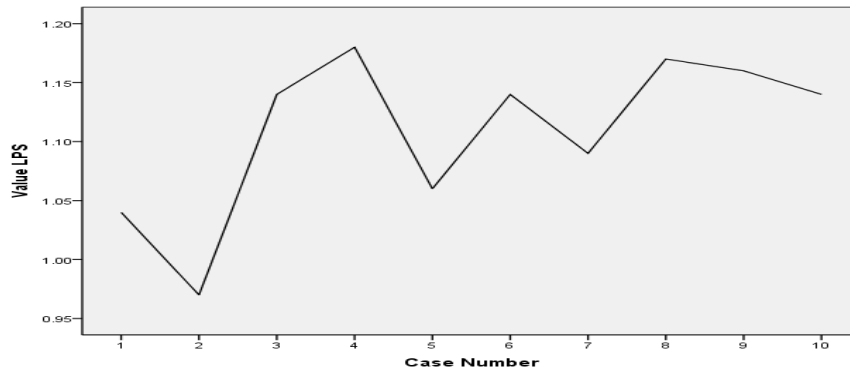

Figure1. *Distribution of learning potential scores among learners with the same non-dynamic score of 30*

Also, regarding the last research questions, the potential for learning among high and low achievers through computerized mediation was investigated. One of the primary objectives within the DA way is the notion that mediation tends to have a greater impact on individuals who have achieved lower levels of success; regardless of whether their underachievement stems from diverse cultural, socio-economic, or educational backgrounds (Peña, Iglesias & Lidz, 2001; Tzuriel & Kaufman, 1999). The last research question addresses this matter by examining whether individuals who have achieved low and high levels of success benefited in different ways and amounts from the hints. A comparison was made using an independent-samples t-test to determine the mean LPSs for high achievers and low achievers. As depicted in Table 5, There was a significant difference in the average scores for learning potential between the two groups (P. <.000). The magnitude of the difference was large. (Eta squared = .9).

Table 5.
Independent Samples T-Test for LPSs on the Tests

	t	df	p	std. error difference
LPS Equal variances assumed	6.44	80	.000	.03

Discussion

As for the first research question, the findings of the study were similar to Barabadi's (2010) study that developed a computerized dynamic test of reading comprehension. In his study, the test takers outscored significantly in their dynamic score. The results of this study were also in line with those of Poehner and Lantolf (2005), though they followed an interactionist approach. In their study, the test takers increased their awareness of the grammatical concepts after they received mediation through DA sessions.

The reason behind the findings obtained from this investigation could be explained as follows: (a) as most of all of the test takers have benefited from the hints in different percentages and manners, DA

researchers claim that static (non-dynamic) tests are incapable of providing a comprehensive picture of the test taker abilities, that is both intramental and intermental, is verified. Whereas static tests can only explain the intramental, self-regulated, and fully internalized capabilities of the learners, dynamic tests have the capacity to assess not only these abilities but also the intramental and other-regulated abilities. From a Vygotskyan standpoint, it can be argued that non-dynamic tests are limited in their ability to capture the learners' Zone of Actual Development (ZAD). Conversely, dynamic tests have the capacity to encompass both the ZAD and the Zone of Proximal Development (ZPD). (b) As Barabadi (2010) claims the improvement of test takers' scores from a non-dynamic to a dynamic test may not be exclusively due to the learners' ZPD. When we examine the scoring files of the test takers and consider the amount and type of hints used, it becomes clear that a considerable number of test takers have answered certain questions after receiving the initial hint, which states, "Your answer is Wrong! Try again." This suggests that factors like demotivation and lack of concentration could be the primary causes of the test takers' incorrect response in their initial try. In other words, while the initial clue is not dependent on the grammatical aspect's nature, it aids and supports the learners in overcoming these factors that could potentially result in a complete loss of points in a static test.

An important issue regarding LPS is the way it can be interpreted, that is what it means when it is said that a learner has a high or low LPS. Kozulin and Garb (2002), who first developed this formula, believe that a high LPS means that a learner's ZPD is almost at the same level as ZAD. In other words, the ability that is aimed for is about to be internalized or self-regulated. However, a low LPS signifies that the ability or the knowledge in question is still far from being self-regulated. In other words, a learner with a low LPS needs more mediation based on hints and prompts in order to find the right answer. In this study, relatively similar results were obtained in that the more learners made use of the hints, the lower their LPS score. Pearson r correlation was used in order to find the kind of relationship between the learners' LPS and the number of mediations they used in the dynamic

test. In the current study, we came up with a correlation of r =-.83 which was significant at P< .01.

The findings of the present study are also congruent with those of Kozulin and Garb (2002) and Barabadi (2010). They both came across a relatively similar conclusion by which LPS can differentiate among the students with the same non-dynamic score. Through investigating the role of computerized dynamic tests of reading comprehension Barabadi (2010) illustrated that LPS can differentiate between the learners with equal non-dynamic scores. Kozulin and Garb (2002) also reported that LPS was able to differentiate between the learners with the same (non-dynamic) pre-test score. In many other DA studies, such as Poehner and Lantolf (2005) and Anton (2009), studies have found that DA can distinguish among individuals with equal scores in non-dynamic tests.

Conclusion and Pedagogical Implications

C-DA represents a significant advancement within the realm of DA since it gives the authority to teachers in order to calculate a considerable number of learners in a dynamic approach concurrently. When computers are able to assume the role of mediator between examiners, the need for teachers in the classroom is greatly diminished. Learners are able to engage with their computers as a means of mediation. Additionally, through the process of monitoring learners' mistakes, C-DA provides the opportunity for both educators and students to analyze and identify their areas for improvement. Later, practitioners can turn their focus of attention to the problematic areas of their learners. C-DA facilitates students' ability to assess and reassess themselves; it prompts them to actively engage in the entire learning and assessment process. Hence, given the existence of C-DA, students are now emancipated from their dependence on teachers for assessment and self-awareness of their progress; they have the ability to assess and reassess their own advancement as frequently as required.

C-DA is a very recent phenomenon within the field of L2 assessment through which instruction and evaluation will be unified so as to increase learners' development. Since DA procedures

consider both latent and emerged capabilities while assessing learners, it appears to be reasonable to suggest the utilization of DA in conjunction with traditional assessments. Henceforth, it is significant for educators to comprehend this reality that the discerning utilization of these two categories of assessment furnishes them with a distinct picture of learners' capabilities; a picture that considers not solely the existing advanced abilities but also the developing ones. Teachers were prevented from the overestimation and underestimation of their students' abilities by utilizing the data and information acquired through the application of DA.

DA procedure like the one conducted in this particular study aids teachers in avoiding the tendency to overestimate or underestimate the capabilities of their students. This is mainly effective when highly sensitive decisions about the individuals' lives are made based on the findings of the test, e.g., a high-stakes established test such as university entrance examination.

The findings of this study have specific utility for teachers, as C-DA offers them insight into both the actual level of performance and the learning potential of their students. Through a comparison of students' non-dynamic scores and their LSP, instructors can develop customized learning strategies to accommodate various learning requirements. To put it differently, two students with equal non-dynamic scores but differing high and low learning potential scores can be treated differently. The student with a high LPS should be provided with more challenging resources and opportunities for autonomous study, as their high LPS indicates a level of internalized ability. Conversely, the student with a low learning potential should be taught learning and information-processing strategies. In other words, they should be taught how to learn, as their abilities are still in a state of being externally regulated. Similarly, the teacher can develop distinct plans for each individual student. That being so, the scoring file presents an elaboration of the score gained and the number of mediations used for each question. Both the teachers and the learners can make use of this information. The teachers can take out those questions that have been answered by the learners with a high number of hints used as the problematic

questions. They can work on these questions in make-up sessions in order to make sure that learners have learnt the grammatical points completely. The learners can also study their scoring profile in order to determine their weaknesses and work on them privately. In sum, this research has a number of significant implications for further study for language teachers, language programmers, language learners, and educational institutions.

Language Teachers

Scoring profiles and LPS play a crucial role in providing teachers with a clear understanding of their students' abilities. Therefore, it is recommended that L2 teachers use C-DA programs to distinguish their learners' strengths and shortcomings and increase individualized learning strategies.

Language Learners

Learners can greatly benefit from the detailed scoring profiles generated by C-DA programs. They can analyze their performance and become aware of their strengths and weaknesses. This will enable them to plan effectively for their future performance.

Educational Institutions

The main objective of educational institutions is to enhance the learning outcomes of their students, such as high achievement rates and admission to high-ranking universities. Therefore, using diagnostic tests to gain a comprehensive understanding of students' abilities, weaknesses, and strengths is crucial in achieving this goal.

Suggestions for Further Research

Considering both the findings and limitations of the present study, the suggestions for further research are as follows: Firstly, the researchers are recommended to design and implement online interventionist CDA procedures of listening and reading skills through formats other than the multiple-choice format in which the inherent problems of multiple-choice tests are resolved. Secondly, as the development of the learners' abilities is at the heart of each C-DA procedure, the researcher proposes designing and implementing

an online C-DA project for Persian language proficiency for those who are interested in learning Persian as a foreign language. So far, we have online interventionist C-DA for languages such as Chinese, French, and English.

References

Ableeva, R. (2008). The effects of dynamic assessment on L2 listening comprehension. *Sociocultural Theory and the Teaching of Second Languages*, 57-86.

Ahmadi, A., & Barabadi, E. (2014). Examining Iranian EFL learners' knowledge of grammar through a computerized dynamic test. *Issues in Language Teaching, 3*(2), 183-161.

Ajideh, P., & Nourdad, N. (2012). The immediate and delayed effect of dynamic assessment on EFL reading ability. *English Language Teaching, 5*(12), 141.

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.

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

Barabadi, E (2010). *Designing computerized dynamic assessment of L2 reading comprehension of Iranian university students and its comparison with static test of L2 reading comprehension.* Unpublished thesis. Ferdowsi University of Mashhad.

Birjandi, P &. Ebadi, S. (2010). *Exploring learners' microgenetic development in L2 dynamic assessment via online web 2.0 technology.* Paper presented in IELTI5, University of Tehran, Tehran, Iran.

Dornyei, Z. (2007). *Research methods in applied linguistics.* Oxford: Oxford University Press.

Jacobs, E. L. (2001). The effects of adding dynamic assessment components to a computerized preschool language-screening test. *Communication Disorders Quarterly, 22* (4), 217–226.

Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL text comprehension. *School Psychology International, 23,* 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. & Aljaafreh, A. (1995). Second language learning in the zone of proximal development: A revolutionary experience. *International Journal of Educational Research, 23*(7), 619-632

Lantolf, J. P. & Poehner, M. E. (2008). Dynamic Assessment. In E. Shohamy (Ed.), *The encyclopedia of language and education* (vol. 7): *Language testing and assessment* (pp. 273-285). Cambridge: Cambridge University Press.

Malmeer, E., & Zoghi, M. (2014). Dynamic assessment of grammar with different age groups. *Theory and Practice in Language Studies, 4*(8), 1707-1713.

Mehri Kamrood, A., Davoudi, M., Ghaniabadi, S., & Amirian, S. M. R. (2021). Diagnosing L2 learners' development through online computerized dynamic assessment. *Computer Assisted Language Learning, 34*(7), 868-897.

Mc Neil, L. (2016). Understanding and addressing the challenges of learning computer-mediated dynamic assessment: A teacher education study. *Language Teaching Research. 17*(3), 323-342.

Messick, S. (1989). Validity. In R. L. Linn (Ed.). *Educational measurement* (pp. 13- 103). New York: Macmillan.

Modarresi, G., & Alavi, S. M. (2014). Designing and validating a test battery of computerized dynamic assessment of grammar. *TELL, 8* (2), 1-29.

Nasiri, M., & Vahid Dastjerdi, H., & Tavakkoli, M. (2013). Dynamic assessment as a helping hand for English teachers: Computerized assessment of transition words or phrases*, Humanising Language Teaching, 3.*

Nazari, E. T. (2009). *TOEFL sample tests (grammar.)* Tehran: Rahnama Press.

Orikasa, M. (2010). *Interactionist dynamic assessment in L2 learning: A case study of tutoring L2 English oral communication.* Retrieved from http://scholarspace.manoa.hawaii.edu/handle/10125/20258

Panahi, P., Birjandi, P., & Azabdaftari, B. (2013). Toward a sociocultural approach to feedback provision in L2 writing classrooms: the alignment of dynamic assessment and teacher error feedback. *Language Testing in Asia, 3*(1), 1-10.

Peña E., & Iglesias A., & Lidz C. S. (2001) Reducing test bias through dynamic assessment of children's world learning ability. *American Journal of Speech-Language Pathology. 10,*138–154.

Pishghadam, R., & Barabadi, E. (2012). Constructing and validating computerized dynamic assessment of l2 reading comprehension. *Iranian Journal of Applied Linguistics (IJAL), 15*(1), 73-95.

Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting second language development.* Berlin: Springer Publishing.

Poehner, M. E. (2009). Group dynamic assessment: Mediation for the L2 classroom. *TESOL Quarterly, 43,* 471-491.

Poehner, M., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research, 9*(3), 233–265.

Purpura, J. (2004). *Assessing grammar.* Cambridge: Cambridge University Press.

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

Rajaeizadeh, Z. & Biria, R., & Kheirzadeh, S. (2015). Instructional efficacy of dynamic assessment on English vocabulary learning of young Iranian EFL learners: The case of near vs. far transcendence tasks. *Journal of Applied Linguistics and Language Research, 2*(8), 155-168.

Sternberg, R. J., & Grigorenko, E. L., (2001). All testing is dynamic testing. *Issues in Education,7*(2), 134.

Swain, M. (2001). Examining dialogue: Another approach to content specification and validating inferences drawn from test scores. *Language Testing, 18*, 275-302.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.