

The Differing Effect of Computerized Dynamic Assessment of L2 Reading Comprehension on High and Low Achievers

Reza Pishghadam

Ferdoswi University of Mashhad, Iran
Email: rpishghadam@yahoo.com

Elyas Barabadi

Ferdoswi University of Mashhad, Iran
Email: elyasbarabadi@yahoo.com

Ali Mehri Kamrood

University of Tehran, Iran
Email: amehri@ut.ac.ir

Abstract—The present study investigates the effectiveness of using a computerized dynamic reading comprehension test (CDRT) on Iranian EFL students with a moderate level of proficiency. Using an interventionist approach, the researchers have developed a software which provides the learners with pre-fabricated strategy-based mediations for each item. The software provides two scores for each learner; a non-dynamic score that represents the learner's ability to answer the test without using the mediations and a dynamic score which represents the learners' ability to answer the test items using mediations in a stepwise fashion. The results revealed that providing mediation in the form of hints contributed significantly to the increase of students' scores, and consequently to the improvement of their text comprehension. Moreover, this study confirmed that low achievers would benefit more than high achievers from a dynamic test of reading comprehension.

Index Terms—dynamic assessment, reading comprehension, computerized mediation and development

I. INTRODUCTION

During the first decade of the twenty first century many L2 researchers have tried to envisage a monistic view of language instruction and assessment which is termed as Dynamic Assessment (DA) (Ableeva, 2008; Anton, 2009; Birjandi & Ebadi, 2010, 2009; Jacobs, 2001; Kozulin & Garb, 2002; Lantolf, 2009; Poehner, 2008, 2007; Summers, 2008). This post-psychometric view of assessment is a direct attack on the traditional psychometric views that support a dualistic view of instruction and assessment. Based on the Sociocultural Theory of mind (SCT) originated from the Russian psychologist Lev Vygotsky's ideas, DA researchers claim that dialectical integration of instruction and assessment into a dynamic activity will bring about successful education. This requires sensitivity to the learners' Zone of Proximal Development (ZPD) during the assessment procedures which is achieved through the mediators' intervention in terms of providing hints, prompts, and feedbacks.

The sensitivity to each individual learner's ZPD has narrowed down the scope of most DA studies in terms of the number of learners and the number of abilities that are dynamically assessed in a single DA procedure. In response to this problem, DA researchers have recently turned to technology (computer software) to take charge of the mediators' responsibilities, so that a greater number of learners and abilities are assessed in a single DA procedure (Lantolf & Poehner, in press; Summers, 2008; Tzuriel & Shamir, 2002; Jacobs, 2001).

The present study which tries to introduce this new strand of assessment is also more in line with new perspectives toward reading instruction and assessment. Afflerbach (2007) for instance, believes that readers' development should be the main consequence of reading assessment. In fact, readers' development is the cognitive component of reading assessment which focuses on skills and strategies used by developing readers; such process-oriented reading assessment allows teachers to assess in the midst of learning.

II. BACKGROUND

A. ZPD; the Theoretical Foundation of DA

According to Vygotsky (1986), the ZPD is "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" (as cited in Aljafreh & Lantolf, 1994 p. 468). The actual level of development defines development "retrospectively", while the ZPD defines development "prospectively". According to Kozulin, Gindis, Ageyev, and Miller (2003, p.17), the notion of ZPD gives three important insights to the issue of dynamic testing:

1. It focuses our attention on those psychological functions of the child that are emerging at a given moment but that have not yet been fully developed;
2. The concept of ZPD introduces assisted performance as a legitimate parameter of assessment procedure;
3. ZPD helps to conceptualize the difference between the level of actual performance and the learning potential of the child.

B. Dynamic vs. Non-dynamic Assessment

At the dynamic assessment Website "DynamicAssessment.com", DA is defined as "an interactive approach to conducting assessments within the domains of psychology, speech/language, or education that focuses on the ability of the learner to respond to intervention." As Kozulin (2003, p.17) claims: "Instead of studying the child's individual performance, dynamic assessment focuses on the difference between performance before and that after the learning or assistance phase." The most important claim of the DA approach is that the human abilities are "malleable and flexible rather than fixed" (Sternberg & Grigorenko, 2002, p.1). Deutsch and Reynolds (2002) view DA as a turn or return to process-based instruction in which mediational teaching serves as a tool for developing cognitive skills and fostering meta-cognitive awareness.

DA is different from non-dynamic assessments (NDA) in that it looks at the assessment from an ontologically and epistemologically different perspective, that is, the integration of instruction and assessment through intervention in order to develop the abilities being assessed Poehner, (2008). He also points out that DA and NDA do not refer to assessment instruments, but to administration procedures, so he claims that any assessment instrument can be used in a dynamic or non-dynamic fashion. According to Poehner (2008), three important features can distinguish between DA and NDA:

1. The view of the abilities underlying the procedures
2. The purpose of conducting the assessment
3. The role of the assessor

The major concern of the DA approach is that it considers abilities to be "malleable and flexible rather than fixed" (Sternberg & Grigorenko, 2002, p.1). While from a DA perspective abilities are not fixed, NDA considers them as stable traits; the goal of DA is to promote development whereas NDA tries to gather information about past achievement and finally the role of the assessor in static tests is just to conduct the assessment but assessor in DA acts as a mediator (Poehner, 2007; Elliot, 2003).

C. DA in L2 Context

Jacobs (2001) explored the effectiveness of a new version of KIDTALK with preschoolers by adding more dynamic assessment features to the test. The results showed that dynamic assessment had higher predictive validity especially for children coming from low socio-economic status and minority children who are culturally and linguistically different from their mainstream.

Lin (2009) conducted an interactive DA study in an EFL context. He focused on children's speaking and listening skills through providing them with a set of pre-formulated supportive hints and mediations during the test. He concludes that administering such a test would provide us with information about students' needs and their potential responses to mediation. He claims that a successful interactive DA program should meet three factors:

1. It should have clear objectives
2. It should include meaningful tasks that are in the learner's ZPD and that accommodate to pre-formulated hints and mediations.
3. It should enjoy an appropriate rating scale an appropriate analytic approach

Pohner and Lantolf (2005) tried to dynamically assess and promote the French learners' speaking ability focusing on the use of two grammatical concepts of *passé composé* and *imparfait*. They also followed an interactionist approach. They concluded that providing mediation would enhance the students' awareness this grammatical concepts which would result in speaking French more accurately. They also claimed that a DA procedure enables the teacher to distinguish the differences between the students who scored the same in a NDA procedure through providing the learners with mediation which results in distinguishing the learners' ZPD.

D. Purpose of the Study

One of the greatest challenges of DA whose aim is to integrate instruction and assessment is to find appropriate mediation or treatment that is effective for a large number of students, (Haywood & Lidz, 2003). To date, nearly all studies done in DA are case studies in which a limited number of participants have taken part, (Lantolf & Pohner, 2008; Ableeva, 2008, Anton, 2009; Birjandi & Ebadi, 2009; Kozulin & Garb, 2002). Based on the above-mentioned post modernistic view of reading assessment, this study ties to develop a computerized dynamic test of reading

comprehension that could test a large number of students' at the same time. This article tries to provide answers for these questions:

1. Does DA procedure affect the EFL reading comprehension of Iranian students with moderate proficiency in English?
2. Do high and low achievers differ in their use of mediation in the form of hints?

III. METHOD

A. Participants

104 university students participated in the study. They were primarily MA students of TEFL. The mean age of the sample was 28 years, ranging from 18 to 44. The participants were randomly selected from different Iranian universities. Persian was their first language and English was their second language. In order to make our sampling fairly homogenous in terms of their level of proficiency, we just included those students whose non-dynamic scores on this test fell one standard deviation below or above the mean and ignored the rest. So we excluded 27 students, while keeping 77 students (27 males and 50 females) who were at the same level of language proficiency.

B. Instrument

The instrument used in this study was CDRT which stands for Computerized Dynamic Reading Test. This software which had already been developed and run as a part of one of the author's thesis reacts to the students' responses automatically through providing feedback and hints that are contingent on ZPD while providing us with two scores (dynamic and non-dynamic). The following are the steps taken in designing the software which are presented in brief.

C. Test Preparation

The most difficult stage of doing the current research was preparing a valid test for this study. In order to find appropriate passages, many versions of TOEFL, IELTS and GRE books were studied by the researchers. The passages were about to lend themselves easily to strategy-based mediation e.g. hints; not to be biased against or in favor of particular students; and finally to have a readability level suitable for Iranian university students majoring in English including both MA and BA students. In view of what was said, the reading passages for this test were selected from the book "preparation course for the TOEFL Test" by Philips (2007). We selected passages from TOEFL iBT since the items and the kinds of reading skills and strategies covered in this book are more in line with C-DA.

Having selected appropriate passages, we then set out to prepare items for each passage. The original items of the TOEFL iBT could not be exactly copied here since they had multiple choice format which is not compatible with dynamic testing in which we are intentionally giving students hints to find the answer. So, we had to design items that had open responses so that giving hints to test takers did not make students guess the right answer. The following are the kinds of items used in this C-DA.

1. Identification items, e.g. identification of a word, phrase or a sentence from the text,
2. Writing the appropriate answer in the blank space e.g. to choose from the text or write of their own,
3. Choosing the appropriate alternatives from among a number of alternatives given in the item. It is similar to multiple choice format but the likely answer is much more than multiple choice, and consequently there is no problem in giving hints,
4. Insertion items. Test takers are asked to put a given sentence somewhere in the passage that best fits and completes the meaning.

The original passages (non-dynamic form), were followed by 13 and 18 items respectively and each was given 30 minutes to be answered. In the dynamic form of the test, the number of items was reduced to 10 for each passage. These ten items roughly mapped on to the skills covered in TOEFL iBT book.

After passages and items were prepared, the next step was to prepare 5 hints for each item. The hints are arranged from the most implicit to the most explicit. Naturally, in the first hint which is the most implicit one, the purpose is just to signal to the test takers that their answer is wrong and in this way, they are given the chance to go back to the item and try it again, and in the last hint test takers are provided with the right answer. These two hints usually take the following format in the test:

Hint 1 → Your answer is wrong, try again.

Hint 5 → The right answer is

Although the first and the last hint of each item remained fixed through the test, the rest of the hints were mainly strategy-based, and their format and composition varied from item to item depending on the skill involved in that item. In fact, these three hints were mainly taken from the "How to answer the question" section of TOEFL, IELTS and GRE books. In this part of these books which directly follow each skill of reading, there are some helpful guidelines for students about how to answer reading comprehension items. In other words, the above mentioned books along with some other books on reading were a rich source of inspiration for those three hints. An example of an item in dynamic test is presented in Appendix A.

D. The Software Preparation

The software package has been designed in such a way that any PC can run it easily. It can be installed properly on any computer provided that it has NET Framework software installed on it. On the opening page of the software, test takers need to type the required information about their names, age and majors in the blank spaces. The next two pages of the software provide test takers with a short and simple description of the software both in English and Persian, and it is up to test takers which description to choose and read. After reading the description, test takers can start the test. By starting the test, the first passage and its first item appear on the screen. They have 20 minutes to read the text. Having studied the passage, test takers can start answering the items. While test takers are answering the items, they have an unimpeded view of the passage. They have four minutes for each item, and if they cannot answer an item within four minutes, the software will move automatically to the next item, and test takers will miss that item. If a student gives a wrong answer to an item, the software will provide him with hints until he gets to the right answer in the fifth hint. When the test is over, a scoring file is created on the desktop. As it is shown in Appendix B, the following information about each test taker is stored in this file.

1. Test taker non-dynamic score. This score is calculated according to the students' first try of each item. In fact, this score is exactly the same as that obtained in traditional tests. To make it comparable with dynamic score of the test, we calculated this score on a scale of 0 to 100 points; five points for each item.

2. Test takers' dynamic score. This score is calculated according to students' use of hints. The number of hints used by each test taker is subtracted from the total number of hints which is 100. The number that is obtained by this subtraction is the dynamic score. For instance, imagine that a student uses two hints for the first ten items of the test; that is, two hints for each of these items. This student's dynamic score is 80 which is calculated by subtracting the number of hints used by him (here 20 hints) from 100. The non-dynamic score of the same student would be 50 because this student has given wrong answer to the first 10 items of the test, and only after receiving hints he was able to get to the right answers.

3. The number of hints used in each item. The software also takes into account those items missed by test takers by marking the letter "M" in front of them. This mark shows that test takers could not answer the items within four minutes.

4. The total time spent on the test.

E. Test Piloting

Our major goal of piloting the test was to collect information on the test usefulness for the purpose of making revisions in the test. It was in this phase of test making that a number of ELT experts were asked to substantiate the content validity of the test. To standardize the test, the researchers administered the test to a pilot group of 10 students who had roughly the same language proficiency level as the participants of the study, but they were not selected for the study. Specifically, this pre-testing was done to get qualitative feedback on the test e.g. their feelings and reactions toward C-DA, and also to achieve the following objectives.

1. Making some modifications in content of test including both the items and hints,

2. Making some modifications in the software package. Having received constructive feedback on general layout of the software initially from some experts in ELT and later on from test takers, a couple of major changes were made in the software as well.

IV. RESULTS

Providing mediation in the form of hints contributed significantly to the increase of students' scores, and consequently to the improvement of their text comprehension. In order to determine the statistical significance of the difference between means on these two sets of scores, t-test was used. As indicated in Table 1 below, the results of the t-test reveal that there is a significant difference between NDA and DA ($t = 3.9$, $p < .05$). It means that the test takers have outperformed in dynamic test.

TABLE 1:
THE INFORMATION REQUIRED FOR CALCULATING T-TEST

Two types of test	mean (SD)	df	t	P
Non-dynamic test	47 (7/90)	152	3.90	
Dynamic test	76 (7/38)			.05

DA seems to be a bigger help to weaker students than stronger ones. One of the main assumptions within the DA procedures is that mediation, will, in general be more effective for low achievers; no matter their low achievement is due to cultural, socio-economic or academic reasons (Tzuriel & Kaufman, 1999; Peña, Iglesias & Lidz, 2001). To test this assumption and to provide answer for the second research question, participants in our study were divided into two subgroups based on their scores on the non-dynamic test: a high achiever subgroup with non-dynamic scores of 50, 55 and 60 and a low achiever subgroup with non-dynamic scores of 35, 40 and 45.

In order to have equal numbers in each subgroup, three participants were excluded from the latter subgroup so in each subgroup remained 37 participants. While the high achiever subgroup could increase its mean score on dynamic test by 25 points, the low achiever subgroup could increase it to 32. That is, the latter subgroup, on average, achieved a bigger increase (e.g.7 points) than the former. This means that individuals' failure in static tests may be due to lack of

opportunities for learning or specific learning difficulties or even cultural differences. So, in line with Tzuriel (2003), we come to this conclusion that DA is especially useful when static tests yield low scores.

V. DISCUSSION

The results of our study indicate that the monistic view of DA, that is, the integration of instruction and assessment can be accounted for via developing software which provides the test takers with pre-planned hints (mediation). The effectiveness of such a test can be described on two planes: first, the students' improvement in EFL reading comprehension; second, providing us with information concerning students' learning potential. The findings of our study are also in line with those of Poehner (2007) and Kozulin and Garb (2002, 2004). Poehner demonstrated that how DA interactions can promote development and provide insights into the learners' functioning. Similarly, Kozulin and Garb indicated that DA is effective in both improving students' reading ability and understanding about their potential for learning. Obtaining information about learners' potential allows us to have a true picture of their abilities. Antón (2009) drew similar conclusion by saying that teachers will misrepresent learners' abilities if they consider only the results of traditional tests.

The researchers believe that there can be other reasons for the learners' improvement in terms of their score, that is, by automatically providing mediation when needed and also by automatically generating each individual's profile, C-DA allows for individuals' self-assessment. This makes test takers more involved in their process of learning; therefore such a test can help students overcome their non-intellective factors such as lack of motivation, fear of failure and anxiety by making the second language assessment more learner-friendly. Indeed, DA according to Haywood and Lidz (2007) is part of idiographic enterprise in which each individual is not compared to others rather comparison is within person not with reference to performance of others.

The second question of this study investigates the performance of high and low achievers to see which group is more likely to make maximum use of mediation. Many DA practitioners (Tzuriel, 2000; Haywood & Lidz, 2007) who follow Feuerstein's Mediated Learning Experience (1980, 1987) and Vygotsky's (1978) concept of ZPD hold the belief that individuals who have not received adequate MLE in the past would benefit more from mediation provided during DA session than those who had rich learning experiences. Thus, in line with the above mentioned authors, we hypothesized that low achievers would make bigger gains on dynamic test and would narrow the gap with their counterparts in high achieving group. The results of our study indicated that low achievers' improvement was about 32% as compared to 25% improvement of high achievers. Although the previously conducted studies (Tzuriel, 2000, 2003; Tzuriel & Kaufman, 1999) on the effectiveness of DA in narrowing the gap between high achievers and low achievers were mainly concerned with students who came from different socio-economic background, this study confirmed that DA in general is especially useful and helpful for low achievers; no matter their poor performance is due to socio-economic factors or different home and school language experiences.

Finally, we hope our work could have made a contribution, however little, to the emergent research on DA especially in relation to second language assessment. While recognizing that preparation of a rigorous dynamic test in an area like second language acquisition is demanding and labor-intensive, the benefits that it offers to our students are attractive enough to attract other researchers to show its potentials. Still much work is needed to be done. One of the limitations of this study concerned the preparation of hints. They were prepared based on some guidelines of some reading books. Further research here is needed to prepare hints in response to the errors students would make on the original non-dynamic test. This would increase the chances of working within students' ZPD. It is also recommended to other researchers to investigate the effect of mediation that is provided by computer and a human mediator simultaneously; human mediator intervenes only when electronically delivering mediation does not work. Given in this way, mediation would be more effective since it can better scaffold those students for whom the fixed mediation provided by computer does not work; this in turn would create the intersubjectivity between the expert (teacher) and the novice (student) that enables DA to further learner's development no matter what is their current status.

REFERENCES

- [1] Ableeva, R. (2008). The effects of dynamic assessment on L2 listening comprehension. In *Sociocultural theory and the teaching of second languages*. J.P. Lantolf and M.E. Poehner (Eds.). London: Equinox.
- [2] Afflerbach, P. (2007). Understanding and using reading assessment. Newark, DE: International Reading Association.
- [3] 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.
- [4] Anton, M. (2009). Dynamic assessment of advanced second language learners. *Foreign Language Annals*. 42. 3,576-598
- [5] 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.
- [6] Birjandi, P & Ebadi, S. (2009). Issues in dynamic assessment. *English Language Teaching*, 2. 4, 188-198.
- [7] Deutsch, R. & Reynolds, Y. (2000). The use of dynamic assessment by educational psychologists in the UK. *Educational Psychology in Practice*, 16.3, 311-331.
- [8] Elliott, J. (2003). Dynamic assessment in educational settings: Realizing potential. *Educational Review*, 55(1), 15-32.
- [9] Haywood, H. C. & Lidz, C. S. (2007). Dynamic assessment in practice: Clinical and educational applications. Cambridge: Cambridge University Press.

- [10] 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.
- [11] Kozulin, A. & Garb, E. (2002). Dynamic assessment of EFL text comprehension of at-risk students. *School Psychology International* 23, 112–127.
- [12] Kozulin, A. & Garb, E. (2004). Dynamic assessment of literacy: English as a third Language. *European Journal of Psychology of Education*, 19.1, 65-77.
- [13] Kozulin, A. & Gindis, B. & Ageyev, V.S. & Miller, S.M. (Eds.). (2003). Vygotsky's educational theory in cultural context. Cambridge: Cambridge University Press.
- [14] Lantolf, J.P. (Ed.). (2000). Sociocultural theory and second language learning. Oxford: Oxford University Press.
- [15] Lantolf, J. P. (2009). Dynamic assessment: The dialectic integration of instruction and assessment. *Language Teaching Journal*, 42.3, 355-368.
- [16] 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*. Cambridge: Cambridge University Press, 273–285
- [17] Lin, Zheng. (2009). Interactive dynamic assessment with Children Learning EFL in Kindergarten. *Early Childhood Edu J.* Springer Science+Business Media LLC
- [18] Peña, E. D., Iglesias, A., & Lidz, C. S. (2001). Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology*, 10, 138-154.
- [19] Phillips, D. (2007). Preparation course for the TOEFL test, iBT. New York, NY: Pearson Education. Inc.
- [20] Poehner, M. E. & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research* 9, 1–33.
- [21] Poehner, M. E. (2008). Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development. Berlin: Springer.
- [22] Poehner, M.E. (2007). Beyond the test: L2 dynamic assessment and the transcendence of mediated learning. *The Modern Language Journal* 91,323–340.
- [23] Sternberg, R.J., & Grigorenko, E.L.(2002). Dynamic testing: The nature and measurement of learning potential. Cambridge: Cambridge University Press.
- [24] Summers, R. (2008). Computer mediated dynamic assessment: toward an online model of dialogic engagement. In *Sociocultural theory and the teaching of second languages*. J.P. Lantolf and M.E. Poehner (Eds.). London: Equinox.
- [25] Tzuriel, D. (2000). Dynamic assessment of young children: educational and intervention perspectives. *Educational Psychology Review*, 12 .4, 385-420.
- [26] Tzuriel, D. (2003). Foundations of dynamic assessment of young children. In A. S. H. Seng, L. K. H. Pou & T. O. Seng (Eds.). *Mediated learning experience with children*.
- [27] Tzuriel, D & Shamir, A. (2002). The effects of mediation in computer assisted dynamic assessment. *Journal of Computer Assisted Learning* 18, 21–32.
- [28] Tzuriel, D. & Kaufman, R. (1999). Mediated learning and cognitive modifiability: dynamic assessment of young Ethiopian immigrant children to Israel. *Journal of Cross-Cultural Psychology* 30, 359-380.
- [29] What is Dynamic Assessment? <http://www.dynamicassessment.com/id2.html> Transferred from Guide to Peabody Library's collection on Dynamic Assessment, http://campusguides.library.vanderbilt.edu/PBDY_DA (accessed 5/4/2011)

Reza Pishghadam is associate professor in TEFL in Ferdowsi university of Mashhad. His research interests are psychology and sociology of language education.

Elyas Barabadi was born in Taybad, Iran on September 1983. He got his BA in English Literature and MA in TEFL from Ferdowsi University of Mashhad, Iran.

Ali Mehri Kamrood was born in Ferdows, Iran on April nineteenth, 1983. He got his BA in English Literature from Ferdowsi University of Mashhad. Now, he is an MA student of TEFL in University of Tehran, Iran.