

RESEARCH

Open Access



Assessing English reading comprehension by Chinese EFL learners in computerized dynamic assessment

Yanfeng Yang^{1*}  and David D. Qian²

* Correspondence:

71295753@qq.com

¹School of Foreign Studies, Lingnan Normal University, No. 29 Cunjin Road, Chikan District, Zhanjiang City 524048, Guangdong Province, China

Full list of author information is available at the end of the article

Abstract

Background: Predicated on Vygotsky's zone of proximal development (ZPD), dynamic assessment (DA) integrates teaching with assessment in a single activity by providing mediation to help learners perform beyond their independent functioning and simultaneously promote learner development. It advantages over traditional psychometric tests in that DA can not only reveal the learners' actual level of development from their independent performance as what the traditional tests did but also their potential level of development from the mediations the learners need during the mediator-learner interactions. However, the number of learners a DA assessor can help is limited due to the time demand of mediator-learner interactions. To overcome this weakness, researchers are turning to computerized dynamic assessment (C-DA).

Methods: The study aimed to find out the learners' difficulties in reading comprehension, the effects of C-DA and the learners' reflections on the C-DA program. Multiple choice questions in the reading comprehension section were adapted from CET4 (College English Test, Band 4). Two groups of students participated in the present study. In the pilot study, a group of students were asked to write down their thinking process after they finished a reading test, from which the students' difficulties in reading comprehension test were identified and mediations were worked out afterwards. The mediation levels and the scoring scale were also determined and put into the C-DA program devised for the main study. Another group of students were asked to do a reading test using C-DA. After the test, a learner profile capturing the students' performances in the test was generated from the program. Finally, a questionnaire was distributed to collect data for the students' reflections on C-DA.

Results: It was found that the learners mainly encountered three difficulties in their reading comprehension test: (1) new words were the biggest barrier; (2) learners felt hard to locate on the relevant sentences in the original passage from the information provided in the question; (3) their inferential abilities are weak. The analysis of the data generated by the C-DA program indicated that C-DA far exceeded the traditional static assessment for its stronger diagnostic capacities and the deeper insights it provided for subsequent teaching. The results of the questionnaire showed that C-DA was greatly welcomed by the students and believed to be particularly helpful in building up confidence for those low achievers as identified in traditional static assessment.

(Continued on next page)

(Continued from previous page)

Conclusions: The major contribution of the present study was to implement DA principles to larger scale study with the use of computer technology and gain valuable insights for future studies. However, limitations remained and future studies should try to address the issues discussed in the present study.

Keywords: Socio-cultural theory, Zone of proximal development, Dynamic assessment, Computerized dynamic assessment

Background

Predicated on Vygotsky's social-cultural theory (SCT) of mind, especially his proposition of the zone of proximal development (ZPD), dynamic assessment (DA) integrates teaching with assessment in a single activity by providing mediation to help learners perform beyond their independent functioning and simultaneously promote learner development. DA goes against traditional psychometric tests which hold a dualistic view of teaching and assessment. It advantages over traditional psychometric tests in that DA can reveal not only the learners' actual level of development from their independent performance as what the traditional tests did but also their potential level of development from the mediations the learners need during the mediator-learner interactions. Since what the learners need from the mediator during DA indicate how far the learners are from their independent functioning, the potential level of development is more significant for the subsequent instruction and learner development than the actual level of development. However, the number of learners a DA assessor can help is limited due to the time demand of mediator-learner interactions. To overcome this weakness, researchers are turning to computerized dynamic assessment (C-DA), i.e., using computer as the mediator, to administer a large-scale study (e.g., Poehner and Lantolf 2013; Tzuriel and Shamir 2002; Pishghadam et al. 2011). Despite valuable insights gained from the few C-DA studies, the effectiveness of C-DA sessions is still not very clear. Using Chinese EFL learners as the participants, the present study made a preliminary attempt to design a computer program to assess their reading comprehension. The main purpose of the study is to find out the effect of C-DA on L2 reading, i.e., whether it is more revealing of learners' reading abilities as compared to traditional testing. In addition, the study also considers the learners' difficulties in reading comprehension and their reflections on the C-DA procedure.

Theoretical background

According to Vygotsky and his colleagues, all forms of higher functioning are mediated during the interactions with the environment (Leont'ev 1981; Poehner 2008). At the initial stage, individuals rely on mediations to complete tasks. Later, these kinds of interactions are internalized and give rise to new cognitive development. In other words, it is the mediated interactions that enable individuals to complete the tasks independently which were once only possible through the help of others. Given this view of cognitive development, Vygotsky (1998) proposed that the assessment of individual's abilities should not only include those functions that have been fully developed through independent problem solving but also incorporate those that are in the process

of developing. To better illustrate this notion, he put forward a famous proposition—ZPD. 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” (Vygotsky 1978, p. 86). Mediations therefore provide remarkable insights in revealing an individual’s potential development. The major task of DA, then, is to observe an individual’s actual level of development as well as gain insights into his/her potential development sensitive to ZPD.

Empirical studies on L2 C-DA

There are two major approaches of DA, interventionist and interactionist DA (Lantolf and Poehner 2004). They differ from each other in the form of mediation. In interventionist DA, the mediation prompts are pre-scripted and arranged hierarchically from implicit to explicit with a numerical value representing their position in the sequence (Davin 2013; Poehner and Lantolf 2013). In interactionist DA, the mediation is flexible and negotiated in mediator-learner dialogue (Poehner and Lantolf 2013). For C-DA, interventionist approach is appropriate since computers serve as the mediators instead of humans. However, sometimes researchers also implemented the interactionist approach as a complementary method in C-DA so as to better diagnose learners’ abilities (e.g., Tzuriel and Shamir 2002). Since the focus of the present study is on C-DA, the following review mainly focus on those studies using the computer as the mediator, especially in the field of second language assessment or second language research.

Although Kozulin and Garb (2002) did not use computers in their experiment, their study influenced C-DA studies. They employed an interventionist sandwich format, in which mediation was conducted between the pre-test and the post-test, with a group of 23 adult EFL learners at risk who were immigrants to Israel and failed the high school English test. The purpose of the study was to test the feasibility of DA in curriculum-based areas such as EFL. The pre-test was a standard EFL placement test adapted by eliminating the items on vocabulary recognition, and producing for the completion of those items rely largely on prior knowledge. In this way, the aim of the test was to assess the cognitive strategies in the EFL reading rather than rote memory. Following the pre-test, a detailed guideline was designed and the learners were asked to correct their pre-tests on their own. The material used in the mediation phase was the pre-test, and two phases were involved. The first one was the manipulation of the items in the pre-test on lexical, grammatical, and structural knowledge, after which an “information page” providing information needed for the learners to complete this part was distributed to the learners for take-home study. The second phase involved four texts of increasing difficulty and the reading comprehension questions devised to test the learner’s ability in using the text structure, cohesion devices, and background knowledge to extract meaning. After that, a post-test comparable to the pre-test was administered, and the changes between the scores of the two tests revealed the effects of DA conducted in the mediation phase. Results show that the students’ performance increased considerably in the post-test, indicating that they indeed benefited from the mediation. As noted by the authors, the pre-test scores could only reveal “the learners’ actual performance but not their learning potential” (p. 120).

Kozulin and Garb’s study is limited in that it did not provide a detailed illustration of the mediation phase such as specific examples and protocols (Poehner 2005). However,

the study is significant particularly in its devising of the learning potential score (LPS) to capture the learner's potential level of development. The following is the formula devised.

$$\text{LPS} = (\text{S post} - \text{S pre})/\text{Max S} + \text{S post}/\text{Max S} = (2 \text{S post} - \text{S pre})/\text{Max S}$$

(Kozulin and Garb 2002, p. 121)

S pre and S post were the scores of the pre-test and post-test, respectively, while Max S referred to the maximum score, i.e., the full score of the test. LPS took into account not only the pre- and post-test scores but also the maximum score. The range of the LPS in Kozulin and Garb's (ibid) study was from .47 to 1.21, based on which the authors divided the students into three groups: high learning potential group ($\text{LPS} \geq 1$), low learning potential group ($\text{LPS} < .71$), and intermediate learning potential group ($.79 \geq \text{LPS} \leq .88$). In terms of the validity of the LPS, no further studies can be found by Kozulin and Garb. However, the formula is interesting, and Poehner and Lantolf (2013) employed it in computerized dynamic assessment to further test its feasibility in predicting learning potential.

Tzuriel and Shamir's (2002) study is crucial in C-DA although it belongs to the domain of cognitive psychology. The study examined the effects of C-DA on young children's seriation thinking abilities as compared to DA with an examiner. For the C-DA group, the interventionist approach was adopted and the hints were provided from implicit to explicit by a multimedia program if the children's responses were not correct. In one sub-group of C-DA, teachers were present and provided help tailored to the children's needs when the computer program was not effective during the administration, which was similar to interactionist DA. Results showed that intervention through C-DA supplemented with a teacher's help was the most effective in bringing about cognitive changes, followed by mediation provided exclusively by the computer, and mediation with only an examiner turned out to be the least effective.

Enlightened by Tzuriel and Shamir (ibid), Poehner and Lantolf (2013) took an initial step to capture L2 development using the interventionist DA in a computerized setting. The participants were second language learners of Chinese and French. The tests consisted of L2 reading and listening comprehension items in multiple choice format. Each test incorporated non-transfer and transfer items, the difference of which was that the latter were more difficult than the former. Transfer items were provided after non-transfer ones and aimed to find out to what extent the learners could integrate the knowledge they gained from the previous non-transfer tests when the context changed, i.e., in the more challenging transfer test. Before C-DA, the items were adapted by adding one distractor to the original four in order to provide more chances for the learners to learn in the C-DA process. To establish the appropriate levels of mediation, a pilot study was conducted for each item in a one-on-one interaction between a mediator and several individual learners. The mediations that proved to be effective in the pilot were then written into the C-DA program.

The scoring system was different from the traditional test for which either full score or 0 score could be gained. In C-DA, the student's score decreased as his/her attempts increased for the item. If the student's first attempt was correct, s/he would gain a score of 4 for the item. If a second response was correct, a score of 3 would be earned and so

on until the correct answer was provided by the program at which time no points would be awarded.

What the C-DA program generated were the learner profiles that consisted of a series of numbers recording the learner responses in their attempts for each item. Three scores were generated for the non-transfer items: actual score, mediated score, and LPS. The actual score was the score the student gained in their first attempt for each item, and as such, it was either 4 (correct response) or 0 (incorrect response). It represented the students' actual level of development, i.e., their independent performance without mediation. The mediated score was the summed score gained in the initial attempts and those for the mediated responses. LPS was calculated using Kozulin and Garb's (2002) formula. Another score not generated by the program but mentioned by Poehner and Lantolf (2013) was the gain score. It was calculated by subtracting the actual score from the mediated score, which indicated the extent to which the student had benefited from the mediation.

Statistical analysis of the actual and mediated scores showed that learner performance improved significantly under mediation, which indicated that the mediations provided during C-DA were helpful for the learners. To uncover evidence of learning, Poehner and Lantolf (*ibid*) compared several individuals' actual score, mediated score, and LPS they earned for the non-transfer items with those for the transfer ones and came to a conclusion that they "caution against assuming a direct mapping across all scoring categories" (*ibid*: 335), which may cast doubt on the reliability of the LPS formula proposed by Kozulin and Garb (2002). Significance was also found in the correlation coefficients between LPS and the actual score of the transfer items for both the Chinese reading and listening tests but not for the French reading test. Poehner and Lantolf (2013) concluded that "LPS has promise as a predictor of learning" (p. 336).

In a follow-up report of the same project, Poehner et al. (2014) provided a more fine-grained analysis on what was generated for the non-transfer items in the Chinese reading and listening tests. The procedures for the quantitative analysis in this report were similar to those in Poehner and Lantolf (2013), but the authors expanded the discussion on the results found, especially in explaining and illustrating the factors affecting LPS. As noticed by Poehner et al. (*ibid*), LPS is determined by the actual score, the mediated score, and the maximum score concurrently. In order to have a more complete picture of the learner abilities, one should analyze the actual, mediated scores and LPS collectively rather than focus on one single score. Another significant part of the report not found in the previous publication is that the authors demonstrated what could be gained from the learner profiles generated by the C-DA program and the implications it could provide for future teaching and learning. Through grouping the learners by mediation levels and by language constructs, the authors demonstrated a more fine-grained diagnosis of the learner difficulties, which formed the basis of the suggestions provided for subsequent teaching and learning.

Another study assessing L2 reading in C-DA is Pishghadam et al. (2011). Using the interventionist approach, Pishghadam et al. (2011) investigated the effect of C-DA procedure on EFL reading comprehension by Iranian students at an intermediate level of English proficiency. The comparison between the non-dynamic scores and the dynamic scores reveals that the C-DA procedure was effective in differentiating learners' reading abilities. Moreover, low achievers benefited more than high achievers as categorized

according to their non-dynamic scores. Different from Poehner and Lantolf (2013), the format adopted by Pishghadam et al. (2011) were open-ended items.

Whereas the studies mentioned above carried out the C-DA procedure in one single session to assess reading comprehension, Teo (2012) used C-DA as a teaching method to promote Taiwanese college EFL students' inferential reading skills. The study took a "sandwich" format of interventionist DA, i.e., the mediation phase through C-DA was sandwiched between traditional pre-test and post-test. The whole study lasted for 10 weeks and the mediation phase 8 weeks. For each mediation session, three passages followed by one inferential question for each were provided and mediations were offered from implicit to explicit if the responses were not correct. Similar to most of the C-DA studies, the items were in multiple choice format. Results indicated that students made significant progress through the intervention of C-DA as their scores in the post-test outperformed those in the pre-test. In addition, qualitative analysis of students' written reflections in their working portfolio showed that C-DA procedure also promoted students' metacognitive reading strategies in making inferences. However, as noted by the author, the main limitation of this research was its within-group design, i.e., one targeted group throughout the study, which made it less confident to determine whether the improvement was caused by the frequent practices or the C-DA intervention. Therefore, future studies should take a between-group design in order to find convincing evidence to support the effects of C-DA on promoting reading comprehension abilities.

From the above scant literature on C-DA and second language assessment or second language research, more studies are needed in this line. As noted by Poehner and Lantolf (2013), "we are in the beginning stages of what promises to be a lengthy and challenging process" (p. 337). Inspired by the previous studies, the present study attempts to apply C-DA to Chinese EFL learners. The main purpose is to find out the advantages of C-DA over non-dynamic assessment and the students' feedbacks on the procedure. Since reading comprehension is the subject matter, the students' difficulties in English reading comprehension test will also be revealed.

Methods

The present study was conducted to answer the following research questions.

1. What difficulties do the students encounter in their English reading comprehension test?
2. Is the C-DA procedure more revealing in diagnosing learners' abilities in L2 reading comprehension than the non-dynamic assessment?
3. What is the students' feedback on the C-DA procedure?

Two studies were involved: the pilot study with one group of participants (Group 1) and the main study with another group of participants (Group 2) comparable to Group 1. The pilot study was used to identify the difficulties the students encountered in their English reading comprehension test, which would help the researchers to work out the mediations for the students. Since the two groups of students were comparable, the results of the pilot study were applicable to the students in the main study. In the pilot study, the students were asked to write down their thinking

processes for each item, which were then analyzed to determine whether they had a full understanding of each item. The criteria included the understanding of the item question, the correct location in the key portion of the original passage, a full understanding of the located sentence(s), understanding of the options, and correct choice of the answer. If the students failed to meet any one of the criteria, difficulties were identified for the item. The results of the pilot study were used to answer the first research question.

Based on the students' difficulties found in the pilot study, the mediational prompts were written for the main study and the mediational levels for the prompts were arranged from implicit to explicit down the range. That was, with the increasing demand of the mediations, the prompts became more and more explicit until providing the answer and the explanation in the last prompt. The scoring system followed Poehner and Lantolf (2013). After the test, a learner profile was generated by the program, including the actual score for the students' independent performance, the mediated score showing how much help they needed in the test, LPS calculated using Kozulin and Garb's (2002) formula, the gain score which indicated the extent to which the student had benefited from the mediation and was calculated by subtracting the actual score from the mediated score following Poehner and Lantolf (2013), and the record of the students' choice(s) during their attempts for each item. MS Excel and SPSS were used for statistical analysis, the results of which were used to answer the second research question.

After C-DA, questionnaire was used to collect data about the students' feedbacks. Three questions were asked: (1) students' beliefs of the helpfulness of C-DA and the reasons, (2) whether students like or dislike C-DA and the reasons, and (3) students' suggestions on improving C-DA (see Additional file 1: Appendix C). The results of the questionnaire were used to answer the third research question. The following sections were a detailed description for the research design.

Participants

Two parallel classes participated in the study. Based on their English scores in the Entrance Exam to College, the students' overall English proficiency levels were similar. One class consisting of 52 students (Group 1) participated in the pilot study and another class consisting of 50 (Group 2) in the main study. Since four students logged on to the website twice in the main study, they were excluded from the study. Finally, 46 students (26 male and 20 female) were left for further analysis. They were 20 years old in average and had learned English for about 10 years.¹ All the students were freshmen majoring in industrial design in Lingnan Normal University in mainland China. They came to the university in September, 2013. During the study in the university, they had 2-h English classes every week in regular classrooms without computers for the students and another 2-h English classes once every 2 weeks in the language lab where computers were available for everyone. According to the normal practice of the university, students were encouraged to take CET4 (College English Test, Band 4), a widely recognized national English test for college/university students in mainland China, at the fourth semester,² and around 80% of the students passed the exam according to the statistics calculated based on the students' performance in CET4 in the past 3 years. Since the students were in their second semester, CET4 should be a challenge to most of the participants.

Instruments and procedures

Reading comprehension items in multiple choice format were selected as the focus of the present study because it was the most common format in reading comprehension tests and accounted for the largest proportion in the reading section of CET4. In DA, what is most important is to work out an appropriate test sensitive to the learners' ZPD. That is, the test should be neither too easy nor too difficult. The difficulty level of CET4 is around the participants' ZPD because they are supposed to pass the exam in the near future.

Pilot study

According to Vygotsky, to provide mediations sensitive to learners' ZPD is crucial in DA. In the present study, ZPD should be the group's ZPD rather than the individual's because the latter is difficult to access in large-scale studies. As noted by Poehner (2007), mediating the group "not only moves the group's ZPD forward but also promotes the development of individuals" (p. 338). In order to better diagnose the group's ZPD and identify appropriate mediational prompts, a pilot study was administered (see Additional file 1: Appendix A). It was a paper-and-pencil test consisting of one reading passage followed by five multiple choice questions taken from CET4 administered in June, 2012. Apart from writing down the correct choices, the students were asked to write down their thinking processes in completing the task. From the students' thinking processes, their difficulties in reading comprehension were identified. Since the students could not express themselves freely in English, they could use Chinese to complete the task.

Main study

Several steps were involved in the main study, including the selection and adaptation of the test, preparation of the mediations and setting the scoring scales, preparation of the C-DA program, design of the questionnaire, and administration of the test. Two reading passages consisting of 10 questions (five for each passage) were selected from the past official CET4 administered in December, 2012. Since the C-DA procedure was time-consuming, only three questions out of five were chosen for each reading passage in order to insure that all the students could finish the two reading comprehension texts and the follow-up questionnaire in the language lab. Only the first, the third, and the fifth questions for each passage were chosen considering that the revealing of the answer and the explanation to one question in C-DA might influence the students' selection of the answer to the question immediately following it since the relevant portions of the texts concerned to answering the two questions were close. Finally, the questions numbered 57, 59, 61, 62, 64, and 66 in the original test were selected. Following Poehner and Lantolf (2013), one distracter was added to the traditional four-option MCQs in order to maximize the prompts. The adding distracters were written by the present researchers and checked by a teacher majoring in curriculum and assessment. What should be noted was that the adding of the distracter aimed to provide more chances for the learners to learn in C-DA because the ultimate purpose of C-DA was to promote learner development by integrating teaching and assessment.

In preparation of the mediations, the present researchers made full use of what the students had provided in the pilot study and referred to the detailed explanations of the answer to each question in CET4 published by Enbo Education Publishing Company if

necessary. The meditational hints were strategy-based for different types of questions and arranged hierarchically from implicit to explicit according to the principles of DA.

The scoring scale follows that of Poehner and Lantolf (2013), i.e., the student's score would decrease as his/her attempts increase for the item (also see the "Empirical studies on L2 C-DA" section). The students were allowed to make four attempts at responding to each question. If the student's first attempt was correct, s/he would gain a score of 4 for the item. If a second response was correct, a score of 3 would be earned and so on until the correct answer was provided by the program and no points would be awarded. The following information would be generated from the program: actual score, mediated score, LPS, gain score, and the record of the students' selections through the process. The actual score was the score the student gained in their first attempt for each item (either 4 or 0), representing their independent performance without mediation. The mediated score was calculated by adding the actual score for the independent performance (either 4 or 0 for each item) and the score gained with the help of mediations (3, 2, 1, or 0 for each item). The gain score is calculated by subtracting the actual score from the mediated score, which indicates the extent to which the student has benefited from the mediation (Poehner and Lantolf 2013). LPS was calculated using Kozulin and Garb's (2002) formula.

The C-DA program was written by one of the research team members, a computer programmer. Following a detailed description of the functions intended to be achieved in the C-DA program written by the present researchers, he wrote the coding and tested the program several times before its real use.

To answer the third research question, another questionnaire including three questions was designed (see Additional file 1: Appendix C). The first question asks about the students' belief on the helpfulness of the C-DA in improving their reading abilities, while the second question concerns the students' preference for this kind of testing format. The third question encourages the students to provide their suggestions to improve the C-DA procedure. For the first two questions, the students were asked to write down their choices as well as the reasons for the choices. The students were encouraged to complete the task in Chinese.

The main study was administered in the language lab by another research collaborator on April 14, 2014, 1 month after the pilot study. She was the English teacher of the participants. During the administration of the test, she followed the procedures the first researcher wrote for her and discussed with her prior to the test. Because this form of test was new to the students, a short training to familiarize them with the process was necessary. The English teacher spent around 10 min in describing and presenting a demo of the test to the students (Additional file 1: Appendix B). After that, the students began the test by logging on to the website provided by the teacher with their names and student IDs. The procedures went on smoothly. Most of the students finished the test in 25 min or so. Once the students finished the C-DA tests, the teacher distributed the questionnaires to them. The whole test took about 45 min. After that, the teacher collected all the questionnaires and mailed them to the first researcher for data analysis. The data generated by C-DA were packed by the technician of the language lab and emailed to the first researcher.

Results and discussion

After data collection, MS excel and SPSS were used to process the data for statistical analysis. For the data collected from the questionnaire, they were inputted into excel

file and categorized for the qualitative analysis. The following reports the results of the present study.

Diagnosing learner difficulties in reading comprehension

The results of the pilot study served to answer the first research question. The numbers of the students’ first attempts were found to be basically normally distributed (mean = 2.62, SD = 1.248, maximum = 5, minimum = 0). After careful reading and classification of the students’ thinking processes, the following difficulties were identified as the main obstacles in reading comprehension. Firstly, new vocabularies in the reading text constituted the biggest barrier for most of the students. A few students even could not understand the meaning of the question (7/52). Secondly, some students could not locate the relevant part of the original text from the key information embedded in the question. Thirdly, some students could locate the relevant information in the original text, but they still made an incorrect response because of wrong inference. A few students made inference based on their own common knowledge rather than on the original passage.

Considering the above difficulties and the principles of DA, the following framework was developed for the meditational prompts, including the main contents and their arrangement in the hierarchy. In order to ensure that all the students understand the prompts, they were written in Chinese (see Additional file 1: Appendix B for an example³).

Prompt 1: It was the same for all questions if the initial response was incorrect. The content was “Sorry, the answer is wrong. Please try it again.”

Prompt 2: It would be provided if the second response was not correct. The main purpose of this prompt was to guide the students to locate the relevant paragraph(s) of the text from the key words in the questions.

Prompt 3: It would be provided if the third response was incorrect. This prompt would explain what the question meant and narrowed the search space down to one or several sentences.

Prompt 4: If the fourth try was not correct, the correct answer and the explanations would be offered.

Advantages of C-DA

To answer the second research question, the data generated by C-DA were analyzed. Table 1 is a general description of the actual, mediated, and gain scores of the test.

As Table 1 shows, the mean mediated score is higher than the mean actual score and the standard deviation decreased from the actual to the mediated score. Paired sample *t* test revealed significant difference between the two scores ($t = 12.315$, $df = 45$, $p = .000$, Cohen’s $d = 1.22$), indicating that the students’ performance improved

Table 1 A general description of actual, mediated, and gain scores

	Mean (SD)	Maximum score obtained	Minimum score obtained	Number of learners
Actual score	12.09 (5.56)	24	0	46
Mediated score	17.72 (3.73)	24	6	
Gain score	5.63 (3.10)	12	0	

considerably under the mediation. As noted by Poehner et al. (2014), the aim of C-DA is not to increase the students' scores but to capture the learner's ZPD, i.e., to attempt to diagnose the learner's actual and potential level of development. Since the actual score only tells about the learner independent performance, leaving his/her potentials for future instructions unknown. According to Vygotsky, the learner's responsiveness to mediation can reveal his/her ZPD which is more informative for subsequent instruction. In this way, the mediated scores generated by the C-DA can be used to signal the learners' potentials for future learning. This is one of the advantages of C-DA over the traditional static assessments.

Indeed, what is important in the C-DA is the interpretation of the scores gained from the program and its relevance to the concept of ZPD. In terms of the relationship between actual, mediated, and gain scores, correlational analyses revealed that the mediated score experienced a strongly positive correlation with the actual score ($r = .849, p = .000$), indicating that the learners generally benefited from the mediation, whereas the actual and the gain scores were negatively correlated ($r = -.771, p = .000$). As noted by Poehner et al. (2014), as the learners performed well in their independent problem solving, less room was left for improvement in completing the same task, i.e., the ceiling effect. That is why the higher the actual score, the lower the gain score is. The above findings are consistent with those found in the previous L2 C-DA studies (e.g., Poehner and Lantolf 2013; Poehner et al. 2014; Teo 2012).

Another advantage of C-DA is the information provided by the learner profile, i.e., a set of scores generated from the program for each learner. For illustration, some learners' scores were compared and analyzed (Table 2).

As Table 2 illustrates, students who got identical actual scores did not necessarily map onto the same mediated scores. For example, Learners 1 and 2 did not get any answer correct in their initial responses, i.e., in their independent performance, for which the two learners would be identified as at the same level of development by traditional static assessment. However, through mediation, the two learners performed quite differently. Learner 1 got 6 points and the LPS is .50, falling into the low range of LPS according to Kozulin and Garb's (2002) category, whereas Learner 2 earned 10 points, raising his LPS to .83, towards the upper end of the mid range of LPS ($.79 \leq LPS \leq .88$). In other words, the two learners, who were grouped into the same

Table 2 Some learners' scores in the learner profile

Learner	Actual score	Mediated score	Gain score	LPS
1	0	6	6	0.50
2	0	10	10	0.83
3	4	12	8	0.83
4	12	14	2	0.67
5	12	16	4	0.83
6	12	16	4	0.83
7	12	18	6	1.00
8	12	21	9	1.25
9	20	22	2	1.00
10	24	24	0	1.00

group as indicated in the static assessment, were found considerably different according to their mediated performance. In that way, dynamic assessment is more revealing than the traditional static assessment.

Another point worth mentioning is the interpretation of LPS. Kozulin and Garb’s (2002) intention in devising the formula was to use it to indicate how much support the learner requires to move his/her abilities forward. However, the stakeholders should not depend on one single score to interpret the outcome of C-DA but consider all the scores collectively (Poehner et al., 2014). Generally, the LPS increases with the gain score. However, it is not always the case. For example, Learners 1 and 7 both gained 7 points under mediation, but the LPS for the former is only .5 while the latter amounts to 1.0. The reason is that the actual score, the mediated score, and the maximum score collectively influence the LPS according to the formula by Kozulin and Garb. Similarly, Learners 4 and 9 gained 2 points in the C-DA procedure, respectively, but the latter’s LPS is higher than the former’s ($1.0 > .67$) because Learner 9 got a higher actual score than Learner 4. In terms of the predictive validity of LPS, further studies are needed. One way is to incorporate transfer items after the DA procedure, for which Poehner and Lantolf (2013) have attempted in their study, but the result is not very clear as discussed in the previous chapter. Another way is to conduct research investigating the teaching effect on the different groups of learners categorized by their LPS. If the learners of high LPS outperform those of lower LPS, LPS is valid in predicting learning potential; otherwise, its validity is questioned.

Participants’ feedbacks

Participants are the test takers. To know about their reflections of the test is important in evaluating and improving the test, especially when the test format is new to the test takers such as C-DA. Table 3 is a general description of the students’ choices in the first two questions.

The first question asks about the students’ belief on the helpfulness of the C-DA in improving their reading abilities. Forty-two participants out of 46 (i.e., 91%) thought that the testing format was beneficial in improving their reading abilities, while 4 students held a negative view. The reason why the four students thought the procedure was not helpful is mainly because they could not read the questions first in the procedure, which they thought hindered their pace in completing the task. One student wrote that “To read more is the most important in improving reading abilities [regardless of what kind of format it is].”⁴ For those who were positive towards the C-DA format, the reasons can be classified into five categories—*increase of reading pace and better concentration, usefulness of mediations, vocabulary learning, convenience for reading, and others*. The first two categories dominate students’ reflections, accounting for 93% of

Table 3 Crosstabs of students’ choices in the questionnaire

		Q2 (preference of C-DA)		Total
		A (yes)	B (no)	
Q1 (helpfulness of C-DA)	A (yes)	33	9	42
	B (no)	2	2	4
Total		35	11	46

all the reasons. Forty-eight percent (20/42) of the students reported that the C-DA format, especially the time reminder showing how much time left for each item, can help them better concentrate on the test and increase their reading pace. Forty-five percent (19/42) of the students thought that the mediations are useful facilitators, which, as they further explained, guided them to locate the key information in the original passage where the correct answer can be found, helped them find out the reasons why they made mistakes, motivated them to think actively during the process, and gave them chances to try again. One student wrote that “When I was prompted that I was wrong, I did it again and found some details I missed in the first try. I got it right then. This kind of experience gave me a happy feeling of gains.” Another student explained that “The mediations and the explanations of the correct answer can assist me to clarify my misunderstandings of the text immediately and guide me to think on the right track, through which I think I won’t make the same mistakes next time.” For the other reasons, students indicated that this kind of format was interesting, convenient for reading, and useful for vocabulary learning.

The second question concerns about the students’ preference of the test format. As Table 3 shows, a majority of the participants (35/46, 76%) favored the C-DA procedure. Similar to the results found in question 1, students liked the test format mainly because they could get immediate assistance from the prompts when mistakes were made and the reminder of how much time left helped them be more attentive in the test. One student shared that “The mediations arranged from implicit to explicit are helpful for those students of low English proficiency like me.” Another student wrote that “It is encouraging because I can get points easily, not like the previous reading comprehension tests in which responses result in either full score or zero.” From the students’ reflections, it seems that the C-DA procedure is of great benefits to low achievers as categorized by their non-dynamic scores because it can help them establish confidence in learning. Some students viewed this format from another angle. They reported that the way of using computers to test reading comprehension instead of printed hardcopies was easy to operate and helped protect the environment. They thought this kind of format can trigger their interest in learning English and improve their reading abilities. Interestingly, two students who did not think the procedure was helpful in response to the first question showed their preference for the format because it was environment-friendly and interesting. However, eight students who held a positive attitude towards the first question did not like the procedure simply because they could not read the questions before the passage nor could they mark on the reading text. Two students indicated that they were not used to doing the test through computer.

The third question encourages the students to put forward their suggestions to improve the C-DA procedure. What the students were concerned about mainly fell into five categories: the outlook of the procedure, quantity and arrangement of the mediations, the number of passages, time limit, and the co-occurrence of the passage and the questions. Some students suggested making the window of the C-DA procedure more attractive and reader-friendly, for example, setting a background and augmenting the font of the words. Some advised that the choices for each item should be narrowed down to four instead of five and specific mediations

should be provided other than ask the test taker to try again when the first response was not correct. In terms of the number of passages, a few students thought that more passages and other types of questions should be included. Several students suggested that the time limit should be adapted in accordance with the difficulty level of the reading texts. Corresponding to the reasons why some students did not like the format in question 2, they hoped that the procedure could provide a marker for them to underline the key information during reading.

To sum up, qualitative analysis of the questionnaire shows that a majority of the students favored the C-DA procedure and were positive towards its usefulness in improving their reading abilities. More importantly, the procedure promoted the students' learning interest, especially those low achievers who lack confidence. However, the program needs improvement by making the window more reader-friendly, adding the marking function, and presenting the passage and the questions concurrently.

Conclusions

The present study tried to apply C-DA into L2 English reading comprehension test by Chinese EFL learners. It was found that the learners mainly encountered three difficulties in their reading comprehension test: (1) new vocabularies were the biggest barrier, (2) learners felt it hard to locate the relevant sentences in the original passage from the information provided in the question, and (3) their inferential abilities were weak. The analysis of the data generated by the C-DA program indicated that C-DA far exceeded the traditional static assessment for its stronger diagnostic capacities and the deeper insights it provided for subsequent teaching. The results of the questionnaire showed that C-DA was welcomed by the students and believed to be particularly helpful in building up confidence for those low achievers as identified in traditional static assessment. A majority of the students believed that C-DA should be helpful in promoting their reading comprehension abilities.

The major contribution of the present study is to implement DA principles to a larger scale study. However, limitations remain and future studies should try to address the issues. Firstly, the reading passages and questions are relatively few. Future studies can incorporate more passages and questions in the study. Secondly, transfer items are not included in the present study. According to Vygotsky, transcendence or transfer is an important feature of DA, which unfolds whether or not learners can control a more complex and difficult task that builds upon the tasks in the original assessment (Poehner and Lantolf 2013). Future study should incorporate transfer items to reveal learners' abilities of transcendence in conducting the tasks. Thirdly, the questions in the present study are in multiple choice format. Future study should try to implement the C-DA procedure to the other formats such as open-ended questions. Fourthly, the C-DA procedure is only implemented in one single session in the present study. Future studies can expand the scope to a longer span of time in several sessions and test the effects of C-DA on promoting learner development, which may be of great significance to uncover the evidence of transfer since the learners need time to internalize what they have learned during the process. Finally, between-group design is also needed to test the effect of C-DA in promoting learner development compared with that of the traditional teaching/assessment.

Endnotes

¹Three students did not provide the information of how long they have learned English.

²There are altogether eight semesters during students' 4 years' studies in the university.

³The materials in Additional file 1: Appendix B are for demo, not for the real test.

⁴All the quotations in this section were translated from Chinese to English by the first researcher.

Additional file

Additional file 1: Appendix A: Pilot study. Appendix B: Demo of the C-DA procedure. Appendix C: Questionnaire. (DOCX 17 kb)

Funding

The article was supported by the Youth Research Project of Humanities and Social Sciences Sponsored by the Ministry of Education in China (Project No.: 15YJC740118) and the research project sponsored by Lingnan Normal University (Project No.: YW1503).

Authors' contributions

YY (first author) made substantial contribution to the study and the writing. DDQ (second author) reviewed the manuscript, revised, and made adjustments if necessary. Both authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹School of Foreign Studies, Lingnan Normal University, No. 29 Cunjin Road, Chikan District, Zhanjiang City 524048, Gudangdong Province, China. ²English Department, Hong Kong Polytechnic University, Kowloon, Hong Kong.

Received: 10 March 2017 Accepted: 11 June 2017

Published online: 10 July 2017

References

- Davin, K. J. (2013). Integration of dynamic assessment and instructional conversations to promote development and improve assessment in the language classroom. *Language Teaching Research*, 17, 303–322.
- Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL text comprehension of at-risk students. *School Psychology International*, 23, 112–127.
- Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment: bringing the past into the future. *Journal of Applied Linguistics*, 1, 49–74.
- Leont'ev, A. N. (1981). *Problems of the development of the mind*. Moscow: Progress.
- 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 and Research*, 2(6), 1353–1358.
- Poehner, M. E. (2005). Dynamic assessment among advanced L2 learners of French. Unpublished doctoral dissertation, The Pennsylvania State University, University Park, PA, USA.
- Poehner, M. E. (2007). Beyond the test: L2 dynamic assessment and the transcendence of mediated learning. *The Modern Language Journal*, 91, 323–340.
- Poehner, M. E. (2008). *Dynamic assessment: a Vygotskian approach to understanding and promoting second language development*. Berlin: Springer Publishing.
- 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, 323–342.
- Poehner, M. E., Zhang, J., & Lu, X. (2014). Computerized dynamic assessment (C-DA): diagnosing L2 development according to learner responsiveness to mediation. *Language Testing, Special Issue*, 1–21. doi:10.1177/0265532214560390.
- Teo, A. (2012). Promoting EFL students' inferential reading skills through computerized dynamic assessment. *Language Learning & Technology*, 16(3), 10–20.
- 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 in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1998). The problem of age. In R. W. Rieber (Ed.), *The collected works of L. S. Vygotsky: child psychology* (pp. 187–206). New York: Plenum.