

Beyond Traditional Testing: Exploring the Use of Computerized Dynamic Assessment to Improve EFL Learners' Reading

Adeline Teo

Department of English Language and Literature
Fu Jen Catholic University, Taiwan

Abstract

The present study implemented a computerized dynamic assessment (CDA) program in EFL classes at a university in Taiwan. Unlike a traditional assessment, the CDA program provided learners with pre-programmed mediation during assessment. This study examined the effects of the CDA program in enhancing Taiwanese college EFL students' reading skills. The participants in this study were 137 non-English major students, who were divided into a control and an experimental group. The procedure included a pre-test, CDA intervention phase, and a post-test. The three reading skills investigated in the study were identifying main ideas (FMI), using contextual clues (CC), and making inferences (MI). The research questions included: (1) Is there a significant difference in the participants' reading skills between and within the control and experimental groups before and after the CDA program? and (2) What is the frequency of the computerized mediation usage in the intervention phase? Inferential and descriptive statistics were used to answer the research questions. The findings showed that the experimental group improved more significantly than the control group in FMI, MI, and overall reading skills. However, the CDA program did not have a significant impact on the participants' CC skill in either group. The improvement in the FMI and MI skills is believed to be caused by the design of the computerized mediation. Also, in the FMI and MI tasks, the frequency of mediation usage decreased gradually and consistently throughout the intervention sessions. Furthermore, although the participants' in both groups showed significant progress in their FMI skill, the experimental group improved more due to the CDA intervention.

Keywords: *computerized dynamic assessment, EFL reading, mediation, MLE, ZPD*

Introduction

In a traditional assessment, when a teacher determines how well her students understood the materials they had learned in class, she usually gives them a test within a given time period. Then, she grades the test while the students wait to be informed of the results. Students usually perform this test-taking activity individually. Feuerstein, Feuerstein, and Falik (2010) called this type of test “static assessment” because it assumes that individuals are born with a certain intelligence that remains fixed throughout life. Due to the static nature of the traditional assessment, no interaction between the teacher and the students is encouraged during the assessment process. The teacher is not expected to provide any instructions to the students during the test. The static nature of language assessments, according to Poehner (2008), has caused frustration among many L2 teachers who felt that the assessment is distinct from the objectives of teaching. Lantolf and Poehner (2004) also commented that static language assessment is not sufficient to evaluate a student’s ability anymore because it cannot detect the cause of the learner’s weakness in his ability. Furthermore, static assessments are product-oriented, rather than process-oriented. Lantolf and Poehner thus proposed the use of dynamic assessment (DA) and claimed that without the interaction between the examiner and the test-taker during an assessment, a teacher can easily miss out on the learners’ dynamic and modifiable learning abilities.

Literature Background

Theoretical framework

The central theoretical framework of dynamic assessment (DA) is grounded in the concept of Zone of Proximal Development (ZPD) in Vygotsky’s (1978) Sociocultural Theory, and Feuerstein’s (1979) Mediated Learning Experience (MLE). Vygotsky used ZPD to elaborate his insights on the relationship between development and instruction. He defined ZPD 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, 1978, p.86). Vygotsky claimed that when it comes to promoting one’s learning ability, a novice needs the guidance of a more capable counterpart. The responsibility of this higher-level counterpart is to provide constructive mediations, or scaffolding, to the learner. New cognitive functions and learning abilities originate within this interpersonal interaction, and later they are internalized and transformed to become the student’s inner cognitive processes. Thus, Vygotsky emphasized that instruction/mediation should lead to development because a learner’s internalization of knowledge takes place with the help of mediation provided by a more capable counterpart. This learning ability may serve as a better predictor of the students’ educational needs than the static scores indicated in a static assessment.

The second framework that forms dynamic assessment (DA) is the Mediated Learning Environment (MLE) construct, developed by Feuerstein, Rand and Hoffman (1979). Feuerstein et al. pinpointed the main difference between “static” and “dynamic” assessment by stating that traditional assessment is static in nature because it is conducted under the assumption that learners’ performances are inflexible and unchangeable. The key component in MLE is *mediation*. Feuerstein et al. claimed that individuals have the potential to change and are modifiable if provided with an appropriately-mediated learning environment. Through mediated learning, learners can change the way they think and achieve the purpose of cognitive modifiability. Also, learners will develop efficient thinking skills that are necessary to become

autonomous and independent in their learning process. In a properly-mediated learning environment, a mediator is expected to attend to the learner's responsiveness to the mediations and then modify the mediations according to the learner's needs. The model provided by the mediator then enables the learner to move beyond his current capacities. The learner's cognitive ability will eventually change with the assistance of effective mediations.

Background and previous studies of computerized dynamic assessment (CDA)

A few known CDA studies were conducted in the field of education. One was conducted by Jacobs (2001) who investigated the ability of a CDA program called KIDTALK in identifying potential language disorders in preschoolers from culturally and linguistically diverse backgrounds. Jacobs investigated whether the CDA training and scoring would make a difference in the information gathered about the individual children's performance. Tzuriel and Shamir (2002) designed a CDA procedure for the assessment of kindergarten children's seriation thinking abilities related to children's mathematical skills. Another study was conducted by Guthke and Beckmann (2000). They developed computerized versions of the Leipzig Lerntest (LLT) to provide a series of training tasks to help young learners master the skills in classifying objects based on various characteristics such as the objects' forms, sizes, and colors. If learners failed an item, the program immediately took them to a set of training tasks.

In the ESL/EFL dynamic assessment (DA) discipline, most of the known research has been conducted in a human-to-human and one-on-one interactionist DA approach format (Lantolf & Poehner, 2004; Lin, 2009; Kozulin & Garb, 2002, 2004, Poehner, 2007, 2008; Teo, 2012). Poehner (2008), a proliferate scholar in EFL dynamic assessment research, claimed that research in computerized dynamic assessment (CDA) is scarce and he highly recommended future researchers to explore how technology can be used effectively to promote the implementation of dynamic assessment. In the past, there were some concerns about the use of technology in language assessment, such as learners' unfamiliarity with using computer-delivered assessment, in L2 or EFL language testing (Chalhoub-Deville, 2001). This concern may be taken lightly nowadays because researchers found that since about a decade ago, computers have become increasingly prevalent, not only in language education but also in language assessment (He & Tymms, 2005). For example, Taylor, Jamieson, Eignor, and Kirsch (1998) reported in their study that more than 80% of computer-based TOEFL takers were familiar with the use of computers. They also found that lack of experience in using computers did not show adverse effects on the test results. A study conducted by Sims (2010), which involved 210 Taiwanese freshmen, has further supported Taylor et al.'s study. This study compared the effects of paper and computer-delivered reading exams in Taiwan, and the results indicated that there were no significant differences in the measurements of reading comprehension between the two platforms, except that the participants completed the computer-delivered reading exams faster than the paper-delivered ones. Both methods were proven to be equally valid for assigning English reading comprehension tasks to university students in Taiwan. Sims further recommended teachers, testing services, and educational programs to consider adapting reading tests to a computer platform without being overly concerned that students will be disadvantaged by taking a computerized reading test.

Since the 1990s some researchers have also recommended the use of technology to promote Vygotsky's concept of Zone of Proximal Development (ZPD). For example, Dixon-Krauss (1996) suggested that educators should take advantage of technological tools available to realize Vygotsky's vision of designing lessons in a way that facilitates instruction that is slightly ahead of the learner's development. Crook (1991) also claimed that computers could act much

like a human partner or classroom teacher within the ZPD because technology is capable of making the computerized tool pertinent to the mediation periods associated with internalization. In other words, CDA may be used effectively to compensate for the shortcoming of human DA, and act as the “more competent peers” by providing the learners with mediated feedback to enhance their ZPD.

One study found on using CDA on EFL reading was conducted by Pishghadam, Barabadi, and Kamrood (2011). They used an interventionist approach to develop a C-DA program and investigated its effect on high and low Iranian EFL achievers’ reading comprehension skills. The main shortcoming of this study is its brief and over-simplified mediation, which prevents one from examining closely the significant role of mediation in ZPD as suggested by Vygotsky. Their mediation began by simply indicating whether the participant’s answer was right or wrong. The second level of mediation requested the participant to choose a key word in the question, without instructing them “how” to look for the key word. The third level of mediation pointed one key word to the participant, and the fourth level provided another key word. The fifth/final level of mediation provided the participants with the correct answer directly without further explanation. Overall, what is found lacking in the mediation design is the detailed and step-by-step explanation that could provide the participants with in-depth mediation that they need to move beyond their current level. A recent CDA study was conducted by Teo (2012) who investigated the effects of CDA in EFL learners’ inferential reading skill and their metacognition. The limitation of the research is that it was a small-scale within group study that compared the same learners’ performance before and after receiving computerized mediation. It thus was not possible to determine whether the learners’ improvement was a result of natural progress due to passing of time or the positive impact of the CDA program. Also, Teo’s research only focused on examining one type of reading skill, which was the learners’ inferential reading. The present study attempted to overcome the limitation of the previous CDA studies in EFL research described above and gain more comprehensive insights on EFL reading discipline.

The Present Study

To help fully understand the impact of mediation in CDA, the researcher in the present study created a more detailed CDA program than in Pishghadam, Barabadi, and Kamrood’s study. The mediation adapted the interventionist approach, which means that the mediation was scripted beforehand and progressed gradually from implicit to explicit levels with the intention of using the mediation to help learners move beyond their current level. The purpose of this study is to explore the effects of the CDA program in promoting Taiwanese EFL learners’ reading skills in finding main ideas, using contextual clues, and making inferences. Since many Taiwanese college EFL teachers are challenged by larger class sizes, heavy workloads, and time constraints, having one-on-one mediational DA becomes unrealistic. The present study is thus unique in the way that the CDA program provided learners with pre-programmed *mediation* without having to rely on one-on-one or group human mediation to enhance Taiwanese college EFL learners’ reading skills. It is believed to help solve the large class size and time constraint problems.

Research Questions

Between group

Research question 1: Is there a statistically significant difference in the *overall* pre- and post-test scores between the control and experimental groups?

Research question 2: Is there a statistically significant difference in the pre- and post-test scores for each of the three reading skills (finding main ideas, using contextual clues to predict the meaning of vocabulary words, and making inferences) between the control and experimental groups?

Within group

Research question 3: Is there a statistically significant difference in the *overall* pre- and post-test scores within the experimental group and within the control group?

Research question 4: Is there a statistically significant difference in the pre- and post-test scores for each of the three reading skills (finding main ideas, using contextual clues to predict the meaning of vocabulary words, and making inferences) within the experimental group and within the control group?

Research question 5: What is the frequency of the computerized mediation usage in the intervention phase?

Method

Method designs

The research method design in the study took the form of pre-test–CDA intervention–post-test. This means that all the participants first took a static test as a pre-test where no mediation was provided. Then, in the intervention phase, the computerized mediation was provided for the participants in the experimental group through the computerized dynamic assessment (CDA) program. Meanwhile, the participants in the control group worked on the CDA program reading activity without computerized mediation. Once they submitted their answer to each question, they would be informed whether their answer was correct. Afterward, each participant from both groups took the static post-test. This design allowed the researcher to compare the participants' performance with and without the mediation intervention. The reading passages used throughout the study were adopted from past TOEFL samples.

In the intervention phase, the Viewlet software was used to implement the CDA program. The CDA program, serving the purpose of unifying assessment and instruction as a single development-oriented activity, involved preprogrammed mediation for the participants to use as guidance and support. With the provision of mediation, the CDA program was dynamic and interactive in nature because it allowed the participants to learn from their errors in the process. The mediation of the pre-test–intervention–post-test sequence was a teaching phase intended to strengthen the participants' reading strategies. The computerized mediation was organized in a way that it progressed from implicit to explicit level. Flowcharts of the CDA program for each of the three types of reading skills are shown in Appendixes 1, 2, and 3.

Participants and setting

The study was conducted at a university in Taiwan. The participants were 137 EFL college freshmen who were enrolled in a required two-semester course titled "Freshman English for Non-English Major Students" (FENEMS). Their ages were 18 to 20. FENEMS was held two hours weekly for eighteen weeks each semester. At the beginning of the first semester, the participants were assigned to two different intermediate-level FENEMS classes by the university's General Education Center based on their scores in the English subject of the Taiwanese national college entrance exam. Their scores ranged from 58 to 88 out of 100. Then, based on a random drawing, these two classes were assigned by the researcher into a control group and an experimental group. There were 64 participants in the control group and 73 in the experimental group. None of the participants had experienced CDA prior to the present study.

The CDA program took place in a computer lab at the university where each participant had his own computer and could get access to the Internet to work individually on the CDA program.

Prior to implementation of study

To ensure that appropriate reading passages and questions were selected for the pre-tests, post-tests, and the mediated sessions, the following evaluation process was conducted before the study began.

- (1) Two college reading instructors reviewed and selected appropriate previous years' TOEFL reading passages and reading questions for the study. They compared test specifications and test content, and then chose the reading passages and questions they believed were appropriate reading materials for the participants. When the two reading instructors did not agree with each other, a third reader was involved in making a decision. In the end, 7 reading passages and 21 questions were chosen for the pre-test and the post-test individually. Three reading passages and three questions were selected for the weekly mediated sessions.
- (2) A pilot test was implemented with 30 participants. These pilot test-takers had a similar reading level as the participants of the actual study, but were not selected for the study.
- (3) Afterwards, the researcher computed the correlation between the pre- and post- tests scores. Pearson Product-moment correlation coefficient (r) was used to determine the strength of the relationships between pre- and post-tests variables. The r value was 0.873, which indicated a high correlation between these two tests.
- (4) Cronbach's alpha, a measure of internal consistency, was used to test the reliability of each of the three types of questions in the pre-test and post-test taken by the 30 pilot test-takers mentioned above. Table 1 shows that the 3 types of reading items achieved satisfactory internal consistency.

Table 1. Inter-item correlation

Reading Question Items	Pre-test Alpha	Post-test Alpha
Finding Main Ideas (FMI)	.71	.71
Contextual Clues (CC)	.73	.73
Making Inferences (MI)	.72	.71

- (5) After the reading passages and questions were selected, the same reading instructors evaluated the mediation designed for each reading passage that were used for the mediated sessions. The procedure was the same as in Step 1 described above.
- (6) Prior to the actual study, a pre-test was given to both control and experimental groups to ensure that they were at a similar reading proficiency level. There were 7 reading passages in the pre-test. Each reading passage had 3 questions. Altogether, there were 7 FMI questions, 7 CC questions, and 7 MI questions. After the pre-test, an independent t-test was used to determine whether the scores were significantly different between the two groups. The pre-test results are shown in Table 2. The p -value is larger than 0.05 in the participants' overall test and in each reading skill. No significant differences were found in the overall pre-test scores and their scores in each reading skill. Both groups were at equivalent reading levels in the pre-test.

Table 2. Pre-test overall and each reading skill performance

Group	Overall M (SD)	FMI M (SD)	CC M (SD)	MI M (SD)
Control (N=64)	55.48 (16.67)	14.88 (8.42)	19.66 (8.20)	20.75 (10.49)
Experimental (N=73)	56.97 (18.87)	16.29 (8.89)	19.84 (7.55)	20.74 (10.09)
<i>P</i> -value	0.63	0.34	0.89	0.97

Note. M=mean; SD=standard deviation

Present study's implementation timeline

Week 1

The participants took a pre-test. The procedure was described in Step 6 above.

Week 2 to Week 9

The participants took 8 mediation tests, one weekly. Each mediation test had 3 reading passages. Each passage had 1 question. Altogether there were 1 FMI question, 1CC question, and 1 MI question. So, in the 8-week mediation phase, the participants completed 24 reading passages, and went through 8 mediation treatments for each type of reading skill. There was no time limit for each mediated session because the primary purpose of the mediation stage was to provide the participants with as much in-depth reading strategies intervention as needed.

From Week 2 through Week 9, the participants in the control and experimental groups were also given a "Running Record" (RR) to record their learning. Each participant was asked to report on their use of reading strategies during the reading activities, and supply concrete self-selected evidence of their reading abilities. The RR served two purposes. First, it allowed the participants to get used to thinking about their learning process (metacognition) while working on the CDA program. Secondly, RR helped reduce the chance that the participants might randomly choose an answer from the multiple choices given in the CDA program and then move on to the next question. When the participants were requested to record their thinking processes, they had to slow down and justify their answer each time before selecting their answer from the options provided in the CDA program.

Week 10

The participants took a post-test. The level and format of the post-test was equivalent to the pre-test, which had 7 reading passages with 3 questions in each reading passage. Altogether, there were 7 FMI questions, 7 CC questions, and 7 MI questions.

Mediation Designs

The CDA program consisted of three specific mechanisms of mediation suggested by Aljaafreh & Lantolf (1994). First, the intervention was *graduated*, starting at an implicit level and progressively becoming more specific until the appropriate level is reached. In this way, the mediator will be able to estimate the minimum level of guidance the learners need. Secondly, the mediation was *contingent*. For each reading passage in the CDA program, mediation was presented to the learners when they failed to provide the correct answer. Meanwhile, the computerized mediation ended automatically when the learners found the correct answer to the question. Third, the continuous assessment process was conducted in an *interactive* manner, which allowed the learners to interact with and respond to the preprogrammed computerized intervention.

Also, the CDA program was designed with consideration for usability, flexibility, and dynamism. Each page of the program incorporated effective web design principles. The tool that was used to develop this CDA program is Viewlet software. This software uses Adobe Flash

technology and allows educators to create interactive programs that can be saved as executable files and accessed through a web browser on the Internet. The CDA program stored the participants' responses and recorded the number of incorrect responses and the mediation that were activated. This program provided the dynamic mediation to the participants of the study by allowing them to receive instruction while being assessed. Meanwhile, it allowed the researcher to analyze the reading strategies the participants used while trying to answer the questions.

The CDA program included reading passages, reading comprehension questions, and computerized mediation which followed each reading question. The mediation was presented in a structured manner. For example, the participants first read a reading passage shown on the computer screen. Afterward, they were required to answer the reading comprehension questions. If their answers were incorrect, they would be provided with the preprogrammed mediation in numerical order, namely *Mediations Level 1, 2, 3, and 4*, until they obtained the correct answer to each question. When they answered a question incorrectly, they were given up to three chances to work on the same question. Each successive mediation provided more detailed guidance, progressing from implicit to explicit feedback. If the participants were unable to answer the question correctly after the third attempt, they would move on to a new question. Each participant was allowed to work at his own pace.

The content of the four mediation levels in the CDA program is described in detail below:

Level 1: It provides the *most implicit* mediation, focusing on using a *general* question to guide the learners toward the direction where the answer could possibly be. (e.g., *In the passage, it seems like the author is trying to arouse our interest on a certain issue. What is it?*)

For inferential questions, Level 1 mediation involves helping learners to understand the main idea of the passage if their first attempt to answer the question was incorrect.

Level 2: It provides *more explicit* mediation than the one in Level 1. Hints are narrowed down to guide the learners to focus on certain *paragraphs*, or *sentences* while looking for the correct answer. (e.g., *Pay attention to the 3rd sentence through the last sentence. What is the point the author was trying to make here?*)

Level 3: It provides *very explicit, context-specific* mediation, focusing on ONE *sentence, phrase, or word*. The goal is to pinpoint how the specific information can lead the learners to the correct answer. (e.g., *What does the word "critical" in the 3rd sentence tell you how the author felt about the air pollution problem?*)

Level 4: At this level the *correct answer* is provided followed by a *step-by-step* explanation. When a learner arrives at this stage, it usually shows that he is still far from fully mastering the reading strategy required to understand the concept tested.

Appendixes 1, 2, and 3 show the flowcharts of the CDA procedure. A website <http://tinyurl.com/btwy6s2> was also created to show three samples adapted from the present study.

Results

Between group

Research question 1: Is there a statistically significant difference in the *overall* pre- and post-test scores between the control and experimental groups?

An independent t-test was used to compare the post-test results of the experimental group and of the control group to determine whether there were statistical differences. Table 2 shown earlier indicates that both control and experimental groups had no significant difference in their

reading levels in the pre-test, prior to the intervention phase. However, as shown in Table 3, the post-test p -value was 0.00, which indicates that the experimental group's overall score in the post-test ($M=73.64$, $SD=14.21$) is significantly higher than the control group's overall score ($M=58.28$, $SD=13.64$).

Table 3. Post-test overall and each reading skill performance

Group	Overall M (SD)	FMI M (SD)	CC M (SD)	MI M (SD)
Control (N=64)	58.28 (13.64)	18.64 (7.04)	17.66 (7.68)	21.13 (7.81)
Experimental (N=73)	73.64 (14.21)	24.85 (6.93)	21.93 (8.14)	27.07 (5.60)
<i>P</i> -value	0.00	0.00	0.002	0.00

Note. M=mean; SD=standard deviation

Research question 2: Is there a statistically significant difference in the pre- and post-test scores for each of the three reading skills (FMI, CC, and MI) between the control and experimental groups?

The control and experimental groups had no significant difference in each of their reading skills in the pre-test. However, Table 3 shows that the p -values for the control and experimental groups' post-test scores for each reading skill are less than 0.05. The participants in the experimental group scored significantly higher than the control group in FMI, CC, and MI in the post-test.

Within group

Research question 3: Is there a statistically significant difference in the *overall* pre- and post-test scores within the experimental group and within the control group?

Research question 4: Is there a statistically significant difference in the pre- and post-test scores for each of the three reading skills (FMI, CC, and MI) within the experimental group and within the control group?

A paired samples t -test was used to compare the pre- and post-test results of the experimental group and the control group. Table 4 shows that the p -values for the overall, FMI, and MI are smaller than 0.05. This indicates that the difference between pre- and post-test scores for the experimental group in these three categories is significant. The only exception is the CC skill because the p -value was greater than 0.05.

Table 4. Experimental group's pre and post-test scores

Experimental Group (N=73)	Overall M (SD)	FMI M (SD)	CC M (SD)	MI M (SD)
Pre-test	56.97 (18.87)	16.29 (8.89)	19.84 (7.55)	20.74 (10.09)
Post-test	73.64 ((14.21)	24.85 (6.93)	21.93 (8.14)	27.07 (5.60)
<i>P</i> -value	0.00	0.00	0.09	0.01

Note. M=mean; SD=standard deviation

In terms of the control group's performance, Table 5 shows that the p -values for the overall, CC, and MI were greater than 0.05. This indicates no significant difference in the participants' performance in their pre- and post-tests in these three skills. The exception is the FMI skill, which shows a p -value of 0.009 and indicates a significant improvement in the post-test.

Table 5. Control group's pre and post-test scores

Control Group (N=64)	Overall M (SD)	FMI M (SD)	CC M (SD)	MI M (SD)
-------------------------	-------------------	---------------	--------------	--------------

Pre-test	55.48 (16.67)	14.88 (8.42)	19.66 (8.20)	20.75 (10.49)
Post-test	58.28 (13.64)	18.64 (7.04)	17.66 (7.68)	21.13 (7.81)
<i>P</i> value	0.304	0.009	0.154	0.797

Note. M=mean; SD=standard deviation

Research question 5: What is the frequency of the computerized mediation usage in the intervention phase?

Each time when a participant used mediation during the intervention phase in the study, it was recorded. Then, the total number of participants who used the mediation in each session was converted into a percentage. The percentage figure was calculated based on the number of mediation users divided by the total number of participants in the experimental group, and then multiplied by 100. The procedure was applied to each mediation level for each type of reading skill. The descriptive statistics of the data was examined for the varying frequencies in which the participants benefitted from the mediation when answering the reading questions. Since the four-level mediation provided a wide range of reading strategies, varying from using general guidelines to relying on specific clues, the frequency change in the use of mediation helped to identify the development of the participants' reading skills during the computerized dynamic assessment. Thus, the data were used to compare the impact of the preprogrammed computerized mediation at different levels. Table 6 lists the descriptive data of the number of the participants who used mediation in each session, and the frequency of the mediation usage at each level and session throughout the intervention phase.

Table 6. Frequency of four-level mediation use

	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	Session 7	Session 8
<i>Finding Main Ideas (FMI)</i>								
Level 1	49 (67%)	44 (60%)	46 (63%)	30 (41%)	27 (37%)	25 (34%)	21 (29%)	21 (29%)
Level 2	31 (42%)	25 (34%)	22 (31%)	13 (18%)	18 (25%)	10 (14%)	6 (8%)	1 (1%)
Level 3	16 (22%)	17 (23%)	17 (21%)	10 (14%)	15 (21%)	4 (5%)	3 (4%)	0 (0%)
Level 4	5 (12%)	5 (7%)	6 (8%)	6 (8%)	9 (12%)	3 (4%)	0 (0%)	0 (0%)
<i>Using Contextual Clues (CC)</i>								
Level 1	54 (74%)	30 (41%)	7 (10%)	7 (10%)	22 (30%)	25 (34%)	15 (21%)	5 (7%)
Level 2	42 (58%)	10 (14%)	2 (3%)	2 (3%)	9 (12%)	11 (15%)	7 (10%)	4 (5%)
Level 3	23 (32%)	5 (7%)	2 (3%)	2 (3%)	2 (3%)	3 (4%)	1 (1%)	2 (3%)
Level 4	12 (16%)	2 (3%)	2 (3%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)
<i>Making Inferences (MI)</i>								
Level 1	32 (44%)	32 (44%)	28 (38%)	23 (32%)	23 (32%)	17 (23%)	6 (8%)	3 (4%)
Level 2	22 (30%)	15 (21%)	18 (25%)	17 (23%)	19 (26%)	6 (8%)	0 (0%)	0 (0%)
Level 3	11 (15%)	11 (15%)	7 (10%)	6 (8%)	5 (7%)	5 (7%)	0 (0%)	0 (0%)
Level 4	7 (10%)	7 (10%)	5 (7%)	4 (5%)	4 (5%)	4 (5%)	0 (0%)	0 (0%)

Note. Number of participants was 73.

An analysis based on the data shown in Table 6 is presented below:

- (1) For the FMI reading task, the overall change of frequency in the participants' use of mediation was rather consistent for each mediation level throughout the eight-week intervention phase. The participants relied on the mediation at all levels more frequently during the beginning than in the later sessions. For example, in the first three sessions of FMI

task, 67%, 60%, and 63% of the participants utilized Level 1 mediation respectively; 42%, 34%, and 31% used Level 2 mediation respectively; 22%, 23%, and 21% used Level 3 mediation respectively; and 12%, 7% and 8% used Level 4 mediation respectively. However, in later sessions, as the participants became more independent, the frequency of their use of mediation decreased gradually. Toward the end of the study, as in Sessions 7 and 8, the percentage figures in Levels 2 to 3 mediation had dropped to single digits. For Session 7, only 29% of the participants used Level 1 mediation; 8% used Level 2 mediation; 4% used Level 3 mediation; and 0% used Level 4 mediation. In Session 8, 29% used Level 1 mediation; only 1% used Level 2 mediation, and none used any mediation at all at Levels 3 and 4.

- (2) A similar situation happened to the MI task, where the frequency of mediation use decreased consistently. In the first three sessions, 44% to 38% of the participants employed Level 1 Mediation. Throughout Sessions 4, 5, and 6, there was a gradual drop in the use of mediation. In the last two sessions, only 8% in Session 7 (6 participants) and 4% in Session 8 (3 participants) employed Level 1 mediation. For Mediation Levels 2 to 4, the decrease was also gradual and consistent. The change in usage from Sessions 1 to 6 dropped from 30% to 8% for Level 2 Mediation; 15% to 7% for Level 3; and 10% to 5% for Level 4. None of the participants used any of the Levels 2 to 4 mediation provided in Sessions 7 and 8. The consistently decreasing use of the four-level mediation in FMI and MI tasks showed the participants' decreasing need for the mediation in these two tasks during the latter stage of the study. Meanwhile, as shown in Table 4, they scored significantly higher in the FMI and MI sections of the post-test.
- (3) Unlike the FMI and MI tasks, the CC task showed inconsistent change in the frequency of mediation usage throughout the study. For example, there was a high usage of Levels 1, 2, and 3 mediation in the first session, which was at 74%, 58%, and 32% respectively. In the second session, the frequency decreased drastically to 41% for Level 1, 14% for Level 2, and 7% for Level 3 mediation. In the third and fourth sessions, the frequency dropped noticeably further. However, Level 1 mediation climbed back up to 30%, 34%, and 21% for Sessions 5, 6, and 7 respectively. Furthermore, for Level 2 mediation, although it reached 58% in Session 1, the frequency of mediation use stayed at 3% in Sessions 3 and 4. Then, the figures climbed to double digits in Sessions 5 through 7. Level 3 mediation's frequency was at 32% in Session 1, and dropped to 7% in Session 2 and remained single-digit percentages throughout the remainder of the sessions. Level 4 began with 16% and then dropped to 0% and 1% in subsequent sessions. To conclude, after all the inconsistent highs and lows throughout the eight sessions for the CC task, in the last session, the frequency decreased to single-digits for all four mediation levels.

Discussion

Although both control and experimental groups were at a similar level in all three reading skills in the pre-test, significant differences between these two groups were found in the post-test. The participants in the experimental group scored significantly higher than the control group in their overall performance, as well as in FMI, CC, and MI in the post-test.

When the control group's reading scores in the pre- and post-tests were compared, no significant differences were found in all areas, except for the FMI skill. It can be concluded that practicing reading and answering questions without computerized mediation for eight weeks did not help enhance the participants' CC and MI reading skills. One noteworthy aspect is that the

control group showed improvement in FMI even without the CDA program intervention. However, when the control group's FMI post-test score was compared with the experimental group's FMI post-test score, the latter was significantly higher. This scenario can be reasonably explained that *time* is the factor that caused both the control and experimental groups to improve in their FMI skill. However, with the CDA program intervention, the participants in the experimental group showed significantly more improvement.

When the experimental group's pre- and post-test scores were compared, significant differences were found in the overall, FMI, and MI skills. However, the CC skill showed no significant difference. This means that the CDA program did not have an impact on promoting the participants' CC skill. This conclusion can be further supported by the descriptive data shown in Table 6. The mediation was used in high frequency in Sessions 1 and 2, and it dropped significantly in Sessions 3 and 4. Then, the use increased significantly in Sessions 5, 6, and 7, and decreased again in the last session. It showed that the mediation usage frequency was inconsistent throughout the eight-week intervention sessions. Due to the unpredictable mediation usage frequency, there was no evidence that the mediation could help the participants become more independent in their CC skill. On the contrary, the CDA program proved to have helped expand the participants' ZPD in their FMI and MI skills. This can be seen from the Table 6 data that show the participants became more independent throughout the intervention sessions as the frequency of their mediation use gradually dropped. Meanwhile, their post-test scores increased significantly after the eight-week intervention.

A possible reason causing the CDA program to have no significant effect on the participants' CC skill could be due to the nature of the CC skill and the design of the CDA mediation in promoting the CC skill. Throughout the intervention sessions, the participants were provided with more opportunities to receive mediations in FMI and MI skills than CC skill. According to Harvey and Goudvis (2000), the reading skill in making inferences (MI) is a cognitively-demanding skill which requires learners to read between the lines and make educated guesses on certain outcomes, events, or actions based on their understanding of the reading materials. An EFL students' ability in MI reflects whether he understands the overall main ideas of the reading materials as well as his critical thinking skills. Therefore, the teaching of FMI skill was integrated into the mediation design for the MI skill. As shown in Appendix 3, *Level 1 Mediation* in the design for the MI skill included finding main ideas strategies. What this means is that when a participant answered an MI question incorrectly, he would be asked to identify the main ideas of the reading passage. If he failed to select the correct main idea, he would be provided with the main idea followed by a step-by-step explanation on how to find the main idea in Level 1 Mediation. After the overall main idea was provided, the subsequent mediation would focus on guiding the participants to use the implied messages in the passage to make inferences. As a result, the participants experienced twice the amount of practice in the FMI and the MI tasks than in the CC tasks.

Unlike the mediation designed for the FMI and MI skills, the CC skill in this study was considered an independent skill, which was not associated with mediated strategies used in the FMI nor MI skills. The mediation design focused entirely on guiding the participants to use hints from the surrounding contexts to make educated guesses on the underlined words in the passages. Therefore, the amount of mediation the participants received on promoting their CC skill was less than the other two types of reading skills. One aspect that should be pointed out is that the experimental group's CC skill mean scores in the post-test ($M = 21.93$) was higher than in the pre-test ($M = 19.84$), although the difference is not significant. However, without the CDA

mediation, the control group's CC skill actually digressed in the post-test mean score ($M = 19.66$ in the pre-test, and $M = 17.66$ in the post-test).

Conclusion

Based on the results, the CDA program proved to promote improvement in the participants' reading skills in finding main ideas (FMI) and making inferences (MI). With the support of appropriately-designed computerized mediation, the participants could interact with a human-like "more competent other," which helped them gradually become independent and proficient readers. Also, with the help of technology, the CDA program was able to show a list of attempts the participants made when they were trying to solve the reading questions. The number of attempts the participants made, the corresponding answers they chose, and their responses to the mediation can provide teachers with more diverse and in-depth information about a student's performance at the current level than simply a score shown in a static assessment. Furthermore, the gradual progress from implicit to explicit mediation can help the teacher understand to what extent a learner would need individualized levels of assistance for a certain type of question. As a result, the teacher can modify and fine-tune mediations to accommodate the learner's needs in one-on-one human interactions.

Another advantage of the CDA program is that it allows teachers to "teach" reading strategies to students in a situation where the number of students is a serious concern. It can save teachers time and allow them to instruct all their students while assessing them at the same time. It also leaves teachers extra time to provide face-to-face interactions and mediation for learners with special needs. An effective CDA program is a win-win for both the teachers and the students.

Like any other study, there are some limitations in this current study. First, the chance that the participants might have skimmed through the content and randomly guessed and selected the answers throughout the pre- and post-tests, as well as during the mediation phase cannot be entirely eliminated. Secondly, the present study only involved Taiwanese university students who were randomly divided into the control and experimental groups. More variables such as learners' gender and learning styles should be explored to gain a more complete picture of the effects of the CDA program. In addition, as shown in the findings, the computerized mediation did not have a significant effect on the participants' skills in using contextual clues (CC). The research designs employed in the present study could not identify and provide an absolutely certain explanation as to the ineffectiveness of the CDA program on the CC skill. It is thus highly recommended that future research be implemented to explore and develop effective computerized mediation to promote learners' CC skill in reading.

About the author: Dr. Adeline Teo is an assistant professor in the English Department at Fu Jen Catholic University in Taiwan. She teaches courses including Academic Writing and Teaching EFL Reading and Writing. Her research interest includes ESL/EFL dynamic assessment as well as ESL/EFL reading and writing skills development.

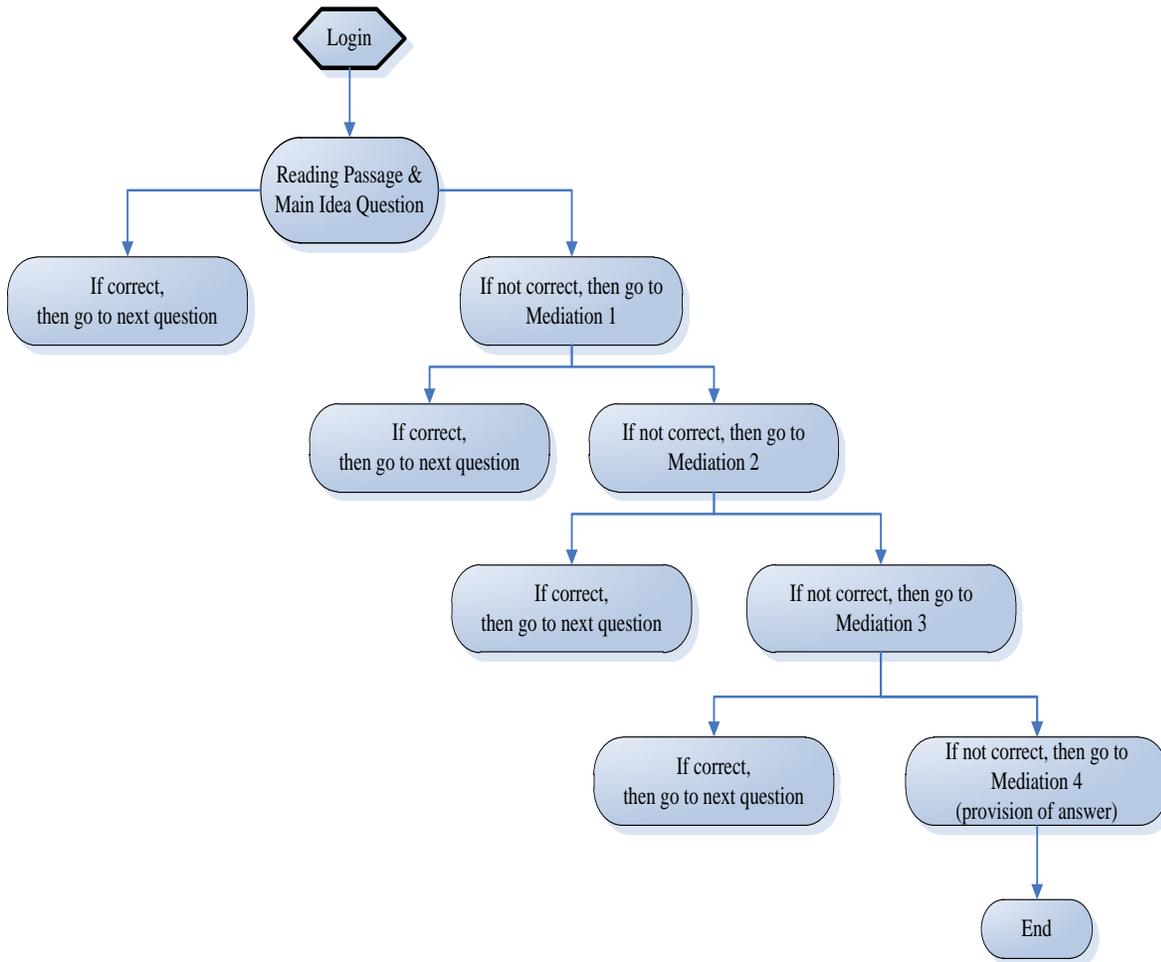
Acknowledgement: This study was funded by the Fu Jen Catholic University Office of Research and Development (Project Grant No. A0102006) and the National Science Council in Taiwan (Project Grant No. NSC 101-2410-H-030-056).

References

- Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 78 (4), 465-483.
- Chalhoub-Deville, M. (2001). Language testing and technology: Past and future. *Language Learning and Technology*, 5(2), 95-98. Retrieved from <http://llt.msu.edu/vol5num2/deville/default.html>
- Crook, C. (1991). Computers in the zone of proximal development: Implications for evaluation. *Computers and Education* 17 (1), 81-91.
- Dixon-Krauss, L. (1996). *Vygotsky in the classroom: Mediated literacy instruction and assessment*. New York: Longman.
- Feuerstein, R., Feuerstein, R. S., & Falik, L. H. (2010). *Beyond smarter: Mediated learning and the brain's capacity for change*. New York: Teachers College Press.
- Feuerstein, R., Rand, Y. & Hoffman, M. B. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques*. Baltimore: University Park Press.
- Guthke, J. & Beckmann, J. F. (2000). The learning test concept and its applications in practice. In C.S. Lidz & J. G. Elliott (Eds.), *Dynamic Assessment: Prevailing models and applications* (pp. 17-69). Amsterdam: Elsevier.
- Harvey, S. & Goudvis, A. (2000). *Strategies that work: Teaching comprehension to enhance understanding*. Maine, Portland: Stenhouse Publishers.
- 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.
- Lantolf, J. P., & Poehner, M. T. (2004). Dynamic assessment: Bringing the past into the future. *Journal of Applied Linguistics*, 1, 49-74.
- 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. (2008). *Dynamic Assessment: A Vygotskian approach to understanding and promoting L2 development issues and implications*. The Pennsylvania State University: Springer.
- Sims, J. M. (2010). Paper- and computer-delivered reading exams in Taiwan. *Studies in English Language and Literature*, 26, 145-157.
- Taylor C., Jamieson, J., Eignor, D., & Kirsch, I. (1998). *The relationship between computer familiarity and performance on computer-based TOEFL test takers* (TOEFL Research Reports, Report 61). Princeton, NJ: Educational Testing Service.
- Teo, A. (2012). Effects of dynamic assessment on college EFL learners' reading skills. *The Journal of Asia TEFL*, 9 (1), 57-94.
- Teo, A. (2012). Promoting EFL students' inferential reading skills through computerized dynamic assessment. *Language Learning and 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: Harvard University Press.

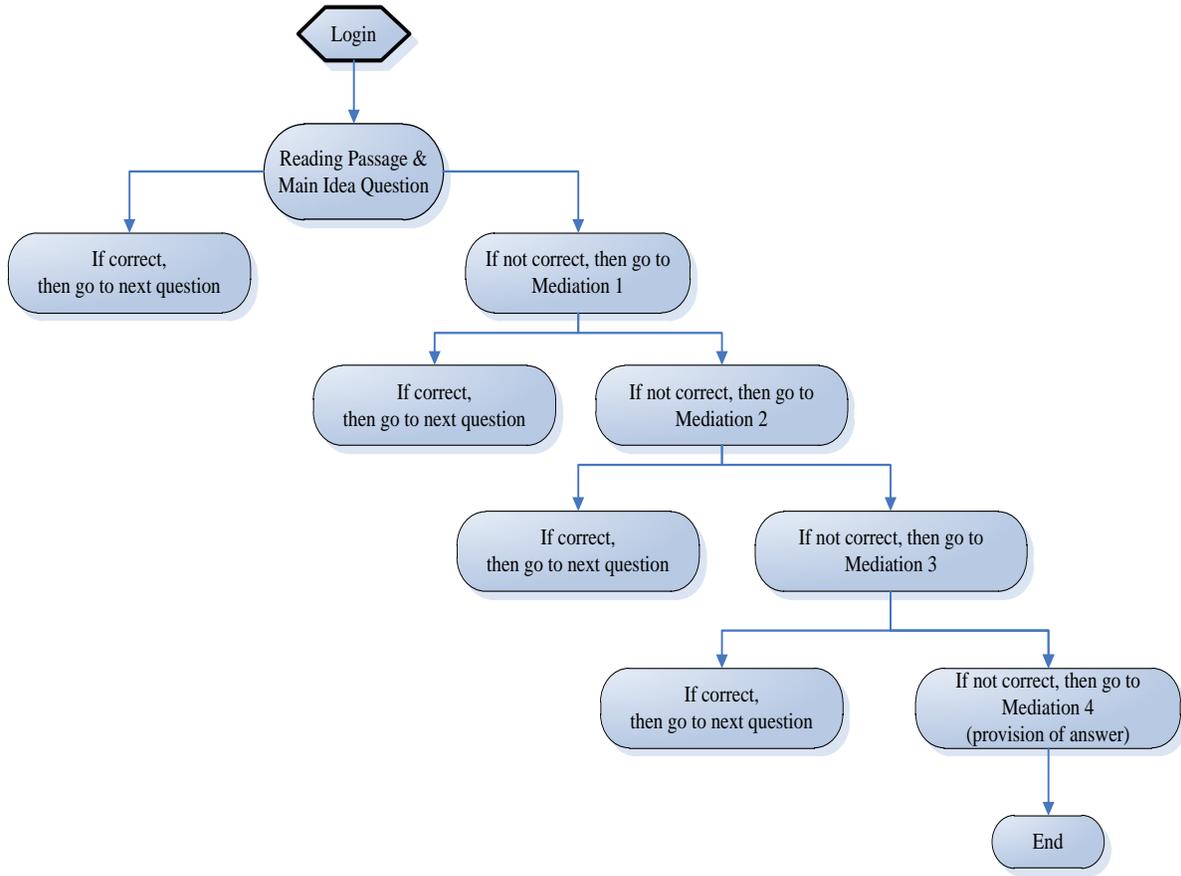
Appendix A

Flowchart for Finding Main Idea Questions



Appendix B

Flowchart for Contextual Clues Questions



Appendix C

Flowchart for Inferential Questions

