International Conference on Current Trends in ELT

The Impacts of Dynamic Assessment and CALL on Critical Reading: An Interventionist Approach

Fatemeh Khonamri\textsuperscript{a}, Mohammad Kazemian Sana‘ati\textsuperscript{b, *}

\textsuperscript{a}University of Mazandaran, Babolsar, Iran
\textsuperscript{b}Iran Language Institute, Rasht, Iran

Abstract

This study investigates the impacts of Dynamic Assessment and Computer Assisted Language Learning on the development of Iranian EFL students' critical reading. Thirty-one male students were chosen in the DACALL and DANonCALL groups, and 30 students were chosen for the Non-DA group. To determine whether CALL procedures would contribute further to the learners' critical reading skills, sixteen subjects from among thirty-one subjects created a special page on Facebook named "Critical Reading Class". The results demonstrated that there were significant differences in scores for DACALL group post-tests, DANonCALL, and Non-DA group reading post-tests.

© 2014 The Authors. Published by Elsevier Ltd. Open access under CC BY-NC-ND license. Selection and peer-review under responsibility of Urmia University, Iran.

Keywords: Dynamic Assessment (DA); CALL; Critical Reading; interventionist; social cultural theory

1. Introduction

Pedagogy and assessment are inextricably interwoven. In this world, hardly anyone refutes the significance of assessment in applied linguistics. At the level of research, testing and pedagogy have emerged as distinct specializations with their own traditions and professional journals and meetings (Pohner2005).

Dynamic Assessment (henceforth DA) refers to the “interaction between an examiner-as-intervener and a learner as- active participant, which seeks to estimate the degree of modifiability of the learner and the means by which positive changes in cognitive functioning can be induced and maintained” (Lidz, 1987, p. 4).

* Mohammad Kazemian. Tel.: 00989111443847;
E-mail address: m_kazemiansanati@yahoo.com
This paper draws on theoretical and empirical research from outside the mainstream paradigm in order to offer an alternative perspective on assessment and its relevance to teaching and learning based on the sociocultural theory of mind developed by L.S. Vygotsky.

Social-cultural theory is about society and culture. The relationship between the social cultural theory and the second language learning is that we learn language socially and culturally which means learning languages social-culturally. The pivotal concept related to the theory is the Zone of Proximal Development (ZPD). Based on the notion of ZPD, therefore, DA seeks not to describe performance constituents but to engage in dialogue with learners in order to arrive at a diagnosis of the sources of performance problems and a prognosis for their remediation based on learner responsiveness during the interaction (Pohner, 2008).

1.1. Statement of the problem

The Static Assessment seems to have failed to assess reading ability in general among poor testees. A superfluity of criticisms is leveled against Static Assessment (henceforth SA). First of all, "traditional group reading assessment measures identified deficits that give EFL poor readers the message that they cannot read without giving them an understanding of what they can do. (Kletzien and Bendar, 1990). Second, Cioffi and Carney (1983) argued that standard assessment procedures are best at evaluating the students' skills knowledge, but insufficient for estimating the students' learning potential and come with little help for identifying the conditions under which the progress can be made. Third, traditional assessment has been criticized for underestimating general ability (Sawanson, 2001) and lacking sensitivity toward so-called disadvantaged students (e.g. Pena, Quinn, and Iglesias, 1992; Utley, Haywood, and Masters, 1992) and students with disabilities (e.g., Lidz, 1987). The study revealed that static assessment could not provide an all-embracing measurement toward students' ability. It is noteworthy that one of the problems of assessment in skills is to approximate the True Score. According to the pilot study that the authors conducted in the EFL classrooms in a certain language Institute (2010), this problem existed in the reading skill in the EFL classrooms in Iran as well. However, this is by no means meant to deny the significance of static assessment.

The present study is going to explore the following research questions:

1. Do CALL-based procedures have any effect on cultivating learners' vocabulary repertoire of post-intervention in the DA group?
2. Do CALL-based procedures have any effect on flourishing learners' reading test scores of post-intervention in the DA group?
3. Is there a positive correlation between post-intervention scores and transfer assessment scores in the DA group?

2. Review of the Literature

The concept of DA stems from both Vygotsky's (1978) learning theory and Feuerstein's (1979) theory of mediated learning experiences. Vygotsky's (1978) conceptualization of a zone of proximal development (ZPD) suggested that learning can be greatly facilitated in interactions between students and a more knowledgeable and experienced person. This fact is realized in a DA approach to teaching and learning. DA literature is rich with different models and approaches which are used to make this interaction possible. Some of these models are briefly sketched here:

Lantolf and Poehner (2004) proposed the interventionist and interactionist approaches to represent two general orientations of DA. The interventionist DA is rooted in Vygotsky's quantitative interpretation of ZPD as a different score. This approach lends itself more to a psychometric orientation. Two well known interventionist approaches are Budoff's "learning potential assessment 'and Brown's graduated prompt' approach (Poehner and Lantolf, 2003). Moreover according to Lantolf and Poehner's (2007), an interventionist approach might use reminders, hints, and leading questions that are weighted differently depending upon their level of explicitness.

Unlike the interactionist DA that follows a strong propensity toward quantification and psychometric analysis, interactionist approaches follow Vygotsky's preference for "qualitative assessment of psychological processes and
As the definition suggests, critical reading is no easy reading and this is where a DA approach may seem really necessary. On the other hand, Computer Assisted Language Learning (CALL) is according to Schmitt and Celce-Murcia (2002, p.9), “one of the more dynamic areas in applied linguistics.” It has been considered as a means of developing various language skills and sub-skills (Marandi, 2011) and thus is worth considering for the combined effect it may have with a dynamic approach which is based on a dialogic interaction between the learners and a more knowledgeable instructor.

3. Methodology

The present study lies within an interdisciplinary framework which includes second language acquisition and Vygotsky's sociocultural theory of cognitive development. The main purpose of this research is to investigate the impacts of DA-based instruction and CALL on the development of Iranian EFL High Intermediate (Upper-Intermediate) students' critical reading. The study applies an interventionist DA approach and is based on quasi-experimental research design.

3.1. Participants

Sixty-one male students were chosen (one standard deviation below and one standard deviation above the mean) out of eighty male students and named High intermediate selected via PET as the experimental group (DA) and thirty students were chosen in a control group(Non-DA).The subjects' participations were voluntary. Furthermore, the experimental group was compartmentalized into two groups, i.e, Dynamic Assessment Facebook group (henceforth DAFG) and Dynamic Assessment Non-Facebook group (henceforth DANFG).

As previously discussed, thirty subjects participated in the DA program. All thirty-one subjects took part in a class called a “Critical Reading Class” in a certain institute. In order to analyze whether a CALL program would contribute further to the learners' critical reading skills, sixteen subjects from among thirty-one subjects created a special page on Facebook called "Critical Reading Class". In other words, this special group was mediated cyberically and non-cyberically and apparently had more privileges than others in the experimental groups. The critical reading classes were held once a week with each session lasting 105 minutes in 10 sessions for both groups. In fact, the class for an experimental group was taught dynamically and for a control group was non-dynamically. Ten sessions (10 weeks) were designated for mediation and treatment. The critical reading class on Facebook was an extra-curricular activity for some of the subjects in the DA class.

Although researchers have elaborated on a number of models and approaches to DA (Budoff, 1987; Guthke, 1993; Carlson and Wiedl, 1992), this study followed Kozulin and Garb's (2001) interventionist approach with some modifications to meet our needs. Regarding the setting and participants in which the research occurred, the interventionist approach was better than the interactionist approach in this study.

A major method of implementing DA in this research was the Graduated Prompt Method, (aka GPA) which was developed by American researchers Campione and Brown.

"The GPA is a variation of interventionist DA in the cake format. This method involves first developing a hierarchy of prompts that are graduated, that is, arranged from most implicit to most explicit. The prompts are standardized with no room for deviation. An especially important feature of GPA, and one which sets it apart from other approaches (Testing-the-Limits), is the inclusion of transfer tasks intended to determine how well students can extend their learning beyond the assessment(ibid p.54)."
The method used in this study was also in line with Kozulin and Garb's (2001) sandwich format, i.e., test – teach – test paradigm. In addition, a follow-up far transfer assessment was conducted to uncover the learner's ability to use (or transfer) newly acquired knowledge to "novel problems" or context.

During the enrichment program, mediation was offered through hints, prompts, questions, suggestions, and explanations.

3.2. Transfer assessment

To establish the extent to which participants could internalize and extend the mediation provided in the course of previous sessions, a transfer session was administered. The transfer session was held precisely after a month of administering the post-test (21st of October 2011).

Each assessment, along with the type of mediation and the properties of the task, is summarized in table 3-1.

Table 3-1. Experiment design: assessment and enrichment program sessions*

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Task Description</th>
<th>Mediation offered</th>
</tr>
</thead>
<tbody>
<tr>
<td>PET</td>
<td>Diagnosing learners’ level of proficiency</td>
<td>None</td>
</tr>
<tr>
<td>Pre-Test for both DA Non-DA groups</td>
<td>Administering SAT via Static Assessment</td>
<td>None</td>
</tr>
<tr>
<td>Enrichment program(10 weeks)</td>
<td>Administering SAT via Static Assessment</td>
<td>None</td>
</tr>
<tr>
<td>Post-Test Non-DA</td>
<td>Administering SAT via Dynamic Assessment</td>
<td>Flexible Intervention with the mediator</td>
</tr>
<tr>
<td>Post-Test DA</td>
<td>Administering a newly developed SAT and TOEFL via Dynamic Assessment</td>
<td>Flexible Intervention with the mediator</td>
</tr>
<tr>
<td>Far Transfer Assessment</td>
<td>Administering a newly developed SAT and TOEFL via Dynamic Assessment</td>
<td>Flexible Intervention with the mediator</td>
</tr>
</tbody>
</table>

*Partially adapted from Poehner (2005).
**Mediation or Flexible interventions involved researcher's leading questions, implicit and explicit hints, prompts, explanations.

3.3. Instrumentation

The materials used in this study and the tasks that learners engaged in include: PET exam as a placement test, Critical Reading Pamphlet used for the mediation and treatment classes, SAT exam, TOEFL exam, and video-recording. It is worth mentioning that all psychometric analyses were done on the SAT exam. The reliability indices were 0.85, 0.87, and 0.86 respectively.

3.4. Procedures

All assessment sessions and the enrichment program (EP) sessions took place within a thirteen –week period (14th of July 2011- 21st of October 2011), excluding ten days of break: September 9th -19th. The assessment sessions were conducted on an individual basis whereas the EP sessions were conducted in a group format.

As previously discussed, the experimental group was classified into two groups, i.e., Dynamic Assessment Facebook (DAFG) and Dynamic Assessment Non-Facebook (DANFG). To determine whether CALL procedures (Computer Assisted Language Learning) would contribute further to the learners' critical reading skills, 16 subjects from among 31 subjects created a special page on Facebook named "Critical Reading Class". More precisely, not only did these students participate in the mediation class, but they also joined in the critical reading section of Facebook.

In the Facebook page, the researchers endeavored to chat with different members of the page about the procedures of critical reading. They inspired the students to be active in the page so that the students could learn and review the lessons and skills from one another. The teachers put a vocabulary puzzle, different SAT clips, and youtube programs. Furthermore, they introduced two sites in the page, www.Sat-preparation.co.il, and www.satprephelp.com to the subjects so that the participants could check their progress in learning new words every single day. It is worth mentioning that whenever the teachers were on line in the
page, they would bombard students with a great deal of vocabulary items. Furthermore, the researchers attempted to give some tips to the learners. For instance, try to learn words with their collocations so that you can make a conjecture whenever you do not know the meanings precisely. For example, deleterious effect, detrimental effect, harmful effect, cacophony noise, cacophony sound, cacophony horn, mellifluous voice. Try to review the strategies of vocabulary learning and do them pragmatically every single day! e.g.: putting the words on the wall.

The researchers put some puzzles for the subjects on the page of Facebook in order to answer them. In fact, the teacher encouraged students to enjoy logomachies.

For instance, Ca------y: a~ noise, a~ horns

Additionally, the teachers used to be on line every night on Facebook and solved students' problems. At the onset of mediation some students found the course sophisticated and complicated and the teachers attempted to inspire them via Facebook.

4. Results

The dependent variable for this study was critical reading. On the other hand, the independent variables were Dynamic Assessment, Static Assessment, and CALL. The findings for raw (Posttest) and gain scores confirmed the main research hypotheses for this study which revealed that those students in the DA groups (CALL and Non-CALL) were expected to produce higher SAT scores than students in the static assessment group.

4.1. The verification of the first null hypothesis

The null hypothesis: CALL-based procedures do not have any effect on cultivating learners' vocabulary repertoire of post-intervention DA group.

To determine whether CALL-based procedures have any effect on cultivating learners' vocabulary repertoire of post-intervention DA group, one-way ANOVA was used for the analysis of the collected data. The SPSS output for the above hypothesis is tabulated in Table 4-1.

<table>
<thead>
<tr>
<th>Result</th>
<th>Levene Statistic</th>
<th>df1</th>
<th>df2</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.584</td>
<td>2</td>
<td>58</td>
<td>.034</td>
</tr>
</tbody>
</table>

Table 4-2. ANOVA

<table>
<thead>
<tr>
<th>Result</th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1616.048</td>
<td>2</td>
<td>808.024</td>
<td>104.821</td>
<td>.000</td>
</tr>
<tr>
<td>Within Groups</td>
<td>447.100</td>
<td>58</td>
<td>7.709</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2063.148</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Subjects were divided into three groups (Dynamic Assessment Facebook Vocabulary group, Dynamic Assessment Non-Facebook Vocabulary group, and Static Assessment (Non-DA) post-test Vocabulary group). There was a statistically significant difference at the P<.05 level in SAT scores for the aforementioned groups [F(2,58)=104.821,P=0.000]. Thus, the null hypothesis was rejected.
Table 4-3. Multiple Comparisons Result Tukey HSD

<table>
<thead>
<tr>
<th>(I) Group</th>
<th>(J) Group</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAGFV</td>
<td>DAGNonFV</td>
<td>3.133*</td>
<td>.998</td>
<td>.007</td>
<td>.73</td>
<td>5.53</td>
</tr>
<tr>
<td></td>
<td>NonDAGPOV</td>
<td>11.567*</td>
<td>.860</td>
<td>.000</td>
<td>9.50</td>
<td>13.63</td>
</tr>
<tr>
<td>DAGNonFV</td>
<td>DAGFV</td>
<td>-3.133*</td>
<td>.998</td>
<td>.007</td>
<td>-5.53</td>
<td>-.73</td>
</tr>
<tr>
<td></td>
<td>NonDAGPOV</td>
<td>8.433*</td>
<td>.878</td>
<td>.000</td>
<td>6.32</td>
<td>10.55</td>
</tr>
<tr>
<td>NonDAGPOV</td>
<td>DAGFV</td>
<td>-11.567*</td>
<td>.860</td>
<td>.000</td>
<td>-13.63</td>
<td>-9.50</td>
</tr>
<tr>
<td></td>
<td>DAGNonFV</td>
<td>-8.433*</td>
<td>.878</td>
<td>.000</td>
<td>-10.55</td>
<td>-6.32</td>
</tr>
</tbody>
</table>

* The mean difference is significant at the 0.05 level.

In order to determine exactly where the differences were in the groups, the post hoc test was also used to illustrate exactly how the performance of the learners differ in terms of their responses to the three tests and whether performance in DAGF differed from performance in DAGNonF and how different performances were in the NonDAGPO. Table 4-3 illustrates the results.

As can be seen from table 4-3, since the paired groups (DAGFV, DAGNonFV), (NonDAGPO, DAGFV), and (NonDAGPOV, DAGNonFV) with the mean differences [3.133, 11.567, and 8.433] were lower than 0.05, it showed that the means of mentioned groups were significant.

4-2-The verification of the second null hypothesis

The null hypothesis: CALL-based procedures do not have any effect on flourishing learners' reading scores of post-intervention DA group.

Table 4-4. Test of Homogeneity of Variances

<table>
<thead>
<tr>
<th>Result</th>
<th>Levene Statistic</th>
<th>df1</th>
<th>df2</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.155</td>
<td>2</td>
<td>58</td>
<td>.322</td>
</tr>
</tbody>
</table>

Subjects were divided into three groups (Dynamic Assessment Facebook Reading group, Dynamic Assessment Non-Facebook Reading group, and Static Assessment (Non-DA) post-test Reading group). There was a statistically significant difference at the P<.05 level in SAT scores for the aforementioned groups [F (2, 58) =144.274, P=0.000]. Thus, the null hypothesis was rejected.

Table 4-5. ANOVA

<table>
<thead>
<tr>
<th>Result</th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1041.449</td>
<td>2</td>
<td>520.725</td>
<td>144.274</td>
<td>.000</td>
</tr>
<tr>
<td>Within Groups</td>
<td>209.338</td>
<td>58</td>
<td>3.609</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1250.787</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Post Hoc Tests

Table 4-6. Homogeneous Subsets

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Subset for alpha = 0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>NonDAGPOR</td>
<td>30</td>
<td>4.67</td>
</tr>
<tr>
<td>DAGNonFR</td>
<td>15</td>
<td>11.47</td>
</tr>
<tr>
<td>DAGFR</td>
<td>16</td>
<td>13.94</td>
</tr>
</tbody>
</table>

Means for groups in homogeneous subsets are displayed.

4.3. The verification of the last null hypothesis

The null hypothesis: There is not a positive correlation between post-intervention DA group scores and transfer assessment scores.

To verify the null hypothesis, the data obtained from post-intervention scores of SAT between DA and transfer assessment scores were subjected to the Pearson product-moment Correlation. The SPSS output for the Pearson Correlation appears in Table. 4-7

Table 4-7. Pearson Correlation (DAGPO and DAGT)

<table>
<thead>
<tr>
<th></th>
<th>2.DAGPO</th>
<th>3.DAGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Correlation</td>
<td>1</td>
<td>.788**</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td></td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.788**</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td></td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>31</td>
<td>31</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (1-tailed).

Preliminary analyses were performed to ensure no violation of the assumptions of normality, linearity and homoscedasticity. There was a strong, positive correlation between the two variables [r=0.788, n=31, P<0.05]. Additionally, due to having a high level of correlation coefficient (0.75), we can use it for an individual prediction.

5. Discussion and Conclusion

The first objective was to explore the effects of CALL and DA on fostering Iranian EFL students' critical reading utilizing an interventionist approach. As elaborated in method, the researchers utilized Facebook as an extra curriculum activity for some subjects. The results indicated that the subjects' performance were statistically more significant than DANonFacebook in the most part. All in all, CALL and DA had an impact on flourishing Iranian EFL students' critical reading utilizing an interventionist approach.

The second objective was to investigate whether subjects with actual level of development, i.e., pretest, showed similar ZPD in the future development, i.e., posttest and far transfer assessment via enrichment program and mediation, and whether posttest and far transfer assessment can predict the subjects' future development. The findings of this study are in line with the results of previous studies in the literature for example Kozulin and Grab (2002), Oskoz (2005), Poehner (2008), Albeeva (2008), Birjandi, et al. (2011), and Naemi and Duvall (2012). Although the aforesaid studies have already been done regarding reading skills and different skills and sub skills, they have paid just scant attention to the reading skills and not on the basis of DA plus CALL and their impact on
critical reading. On the other hand, the researcher could not find any research regarding DA and CALL on critical reading since December 2013. The results of this study indicated that the subjects’ performance improved drastically on the DA posttest and far transfer assessment tests.

Last but not least, the results of the study do suggest that traditional assessment (static assessment), which are the quintessential of most assessment approaches in EFL testing programs tend to underestimate the level of functioning of a learner.

5.1. Limitations of study

Almost every research project suffers from a number of limitations either in its design or in the way it is conducted which restricts its generalizability in one way or another. This research study is no exception and the researchers believe the following points preclude us from drawing strong inferential conclusions from this study:

1. Only one location was selected as the site of the study. Running the study with learners in other educational contexts may yield different findings.
2. All of the participants were only male students. Thus, gender was not taken into account as a variable in this study.

5.3. Pedagogical Implications

Despite the fact that DA only exists in embryonic form of language assessment, some implications can be drawn from the current study. As Kouzlin (2011, P.179) proposed “the development of curriculum–based DA is still lagging significantly behind”.

The implications can be multifold. The first implication concerns instructors who are in charge of a variety of classroom practices including testing sessions. What teachers need to know is the mediation procedures during DA approach. Teachers can employ DA and CALL procedures to lackadaisical students so that they can promote their linguistic competence and language proficiency. The second implication goes to language policy makers and test makers. In fact, they are explicitly dealing with assessment and they are obliged to ameliorate the testing and assessment systems.

References


