Computerized Dynamic Assessment of L2 Reading Comprehension Ability

Nathaniel CARNEY

第二言語読解力のコンピュータによるダイナミック・アセスメント

Nathaniel CARNEY

神戸女学院大学 文学部 英文学科 准教授
連絡先：Nathaniel CARNEY　〒662-8505　西宮市岡田山4-1 神戸女学院大学文学部英文学科
carney@mail.kobe-c.ac.jp

Summary

This article proposes the need for computerized dynamic assessment as a complementary assessment strategy to standardized reading comprehension tests which only offer static portrayals of learners' abilities. First, the differences between dynamic and static tests are examined. Then, an overview of L1 and L2 dynamic reading assessments and computerized dynamic reading assessments to date is provided. Within this overview, particular attention is paid to the types of mediational prompts that have been used in computerized dynamic assessments to date. The need for computerized dynamic assessments with both valid test items and theoretically-based mediational prompts is emphasized. The paper concludes addressing concerns about computerized dynamic assessments and ways to address those concerns.

Keywords: dynamic assessment, L2 reading, mediation, sociocultural theory, CALL

要　旨

この論文では、学習者の能力を静的結果だけで示す標準的読解力テストに加えて、補完的な評価手段としてコンピューターによるダイナミック・アセスメントの必要性を説いている。まず最初に、ダイナミック試験と静的試験の違いを分析する。そして、L1 と L2 のダイナミック・アセスメントの読解力テストと、現在までのコンピューターによるダイナミック・アセスメントの読解力テストについて概要を提示する。この概要の中で特に着目したのは、今までのコンピューターによるダイナミック・アセスメントで使用されてきた指導法である。特に、効果的な試験項目と理論上の指導法の両方を含むコンピューターによるダイナミック・アセスメントの必要性を重視した。この論文では、結論として、コンピューターによるダイナミック・アセスメントに関する重要事項とその解決法を提示する。

キーワード：ダイナミック・アセスメント、第二言語読解力、社会文化理論、指導、CALL

Introduction

Reading assessment, like many other forms of assessment, is often carried out through standardized tests given to massive groups of school-age and tertiary education learners. In adult EFL environments, the most commonly used standardized tests measuring L2 English reading and other competencies are the *TOEFL*, *IELTS*, and *TOEIC*. While such tests are useful and important for placement decisions and determining general reading levels, they are, of course, not perfect. There remains unexplained variance in how a learner performs on these tests and how they actually perform in real life. High-stakes tests can have serious implications for learners' scholastic or vocational trajectory, not to mention learners' self-perception of their own potential for English language learning. Part of the problem with these tests is that they serve as retrospective measures of static ability, not measures of learning potential. Such norm-referenced tests of reading ability are necessary and useful, but when possible they should be complimented by the inclusion of a battery of interactive assessments that measure readers' dynamic potential, not just their static achievement. Computerized dynamic assessment is a method for measuring such potential.

In this paper I propose that a computerized dynamic assessment that gives sufficient attention to the appropriateness of reading texts and test items, the mediational moves offered in response to learners' answers, and the technological concerns raised by computerized tests, can offer a useful complement to scores learners attain through traditional, static, standardized tests. The purpose of this paper is to give an overview of computerized dynamic assessment of reading comprehension in order to offer a design guide for such assessment.

Traditional assessment, dynamic assessment, and computerized dynamic assessment

Dynamic assessment, a Vygotskian-derived concept, includes a broad range of assessments (Haywood & Lidz, 2007; Poehner, 2008), but are defined in essence by "active intervention by examiners and assessment of examinees' response to intervention" (Haywood & Lidz, 2007, p. 1), though the term *mediation* might be preferred to intervention by those drawing more closely from Vygotsky (e.g., Poehner, 2008). Dynamic assessment in practice has often been used in psychology as a means of assessing learners who show retarded mental development in some aspect (Feuerstein, Rand, and Hoffman, 1979).

Dynamic assessment contrasts with what have been called "static tests" (Sternberg & Grigorenko, p. vii), "traditional tests" (Spector, 1992, p. 354), "normative assessment" (Haywood & Lidz, 2007, p. 6), and "non-dynamic assessment" (Poehner, 2008, p. 13). Basically, most major standardized tests of all types are non-dynamic in that they do not involve mediation from the examiner; in fact, they have extremely controlled procedures in which any talking between the

examiner and learner would be seen as potentially biasing the results. Carney and Cioffi (1990, p. 177) cite Fuchs and Fuchs (1986) concerning reading assessments, writing that "traditional conceptualizations of reading assessment suggest a procedure in which an unbiased examiner administers a test instrument in a standardized format, with the entire process in a neutral (often called clinical) setting."

Thus, dynamic assessment differs from the other assessment in how much the examiner is involved in mediating the test with the examinee, but "any assessment can be conducted in a dynamic or non-dynamic fashion" (Poehner, 2008, p. 13). Indeed, recently, methods and procedures for conducting dynamic assessment and dynamic testing (Sternberg & Grigorenko, 2002) have been used and promoted in a broad variety of contexts, including foreign language learning in general (Davin, 2013; Lantolf & Poehner, 1994; Ozkoz, 2005; Schneider & Ganschow, 2000) and both L1 reading (Blachowicz, 1999; Bridges & Catts, 2011; Carney & Cioffi, 1990, 1992; Cioffi & Carney, 1983; Dörfler, Golke, & Artelt, 2009; Fuchs, Compton, Fuchs, Bouton, & Caffrey, 2011; Guterman, 2002; Larson & Nippold, 2007; Spector, 1992) and L2 or L3 reading (Ajideh, Farrokhi, & Nourdad, 2012; Kozulin & Garb, 2002, 2004; Nourdad & Ajideh, 2012; Poehner & Lantolf, 2013; Shabani, 2013; Teo, 2012). Computerized dynamic assessment, specifically, has been seen as especially feasible because it can meet psychometric standards (Embretson, 2000) or because it allows the assessor to mediate the learning of more learners than non-computerized dynamic assessment (Lantolf & Poehner, 2011; Poehner, 2008).

Assessing reading comprehension

A first step in designing a computerized dynamic assessment of reading comprehension is understanding reading comprehension itself. Paris's (2007) description of reading comprehension implies the challenge presented to reading comprehension assessment professionals:

> To construct meaning, readers must decode words fluently, understand vocabulary, make inferences, and relate the ideas in text to their prior knowledge and experiences. These skills vary with age, experience, instruction, context, and motivation so both the processes and the products of reading comprehension are constructive, multidimensional, developmental, and variable. Thus, reading comprehension is difficult to define simply and measure neatly. (p. 1)

This depiction of reading comprehension's complexity is neither new nor uncommon. Nevertheless, reading comprehension assessment is a common feature of standardized adult L2 English language proficiency tests, and common in many commercially produced reading-focused textbooks and materials for adolescent and adult second and foreign language learners. Reading comprehension sections on standardized tests are often quite simple in appearance, e.g., a text

followed by multiple choice questions about the text. However, large standardized tests go through rigorous validation procedures and item analysis, and efforts are made to determine what factors in texts are being tested by a given item, or from another view, what factors in a text increase or decrease its comprehensibility. Thus, for any reading assessment, including dynamic assessments of reading, determining appropriate texts and items is an essential step. Various factors are known to affect the difficulty of reading texts, including lexical level, topic, and discourse style (Grabe, 2009). Therefore, to select or design appropriate reading assessment texts, these factors should be considered.

Vocabulary level, or lexical complexity, is regarded as the most important predictor of a reading texts' difficulty (Alderson, 2000, p. 73). Schmitt, Jiang, and Grabe (2011) found a relatively linear relationship between vocabulary knowledge and reading comprehension in a large study of L2 English readers of a variety of L1 backgrounds. When designing reading text items, vocabulary profiles of English texts can easily be found with the Range program (Heatley, Nation, & Coxhead, 2002) or Tom Cobb's online VocabProfiler (http://www.lextutor.ca/vp/eng/). Of course, vocabulary knowledge alone will not suffice, as even readers with 100% vocabulary knowledge do not always comprehend 100% of the text (Schmitt, Jiang, & Grabe, 2011). Text topics should be carefully selected — if they involve knowledge the readers do not have, they will be more difficult (Alderson, 2000). Finally, discourse styles are important. Readers who can identify the organizational pattern and transitions in a reading are more likely to comprehend it (Alderson, 2000).

Text construction is not the focus of this article and the discussion has thus been brief, but it cannot be overstated that before a useful dynamic assessment is created, the items should undergo text analysis, item analysis, and validation with learners similar to those who will be later dynamically assessed.

Helping learners comprehend: Metacognitive support

Dynamic assessment involves assessors actively participating in the assessment in order to assess and at the same time further develop examinees' cognitive reading abilities through mediation. So, an important question is, what type of mediation might help learners' develop their reading comprehension abilities? Metacognition has been found to have an important role in reading comprehension. Concerning L1 English reading, the National Reading Panel (2000) found that "comprehension instruction can effectively motivate and teach readers to learn and to use comprehension strategies that benefit the reader" (p. 4-6), and teaching metacognitive strategies continues to receive support in more recent literature (e.g., Afflerbach, Pearson, & Paris, 2008; Ahmadi, Ismail, & Abdullah, 2013; Michalsky, Mevarech, & Haibi, 2009). Given this body of research, it seems reasonable to believe that certain metacognitive interventions could serve as mediation for learners engaged in a computerized dynamic assessment. Nevertheless, which

mediational moves are appropriate? Dorfler, Golke, and Artelt (2009) suggest that "the crucial point within dynamic assessments is to design the best possible guidance within a certain dynamic test in order to exploit individuals' full potential" (p. 77). Finding the "best possible guidance" is without question a process in itself that will require both research and pilot testing. This might be the most critical aspect in the design of a computerized dynamic assessment. That is, even if good texts and test items are designed, without proper mediation, the computerized dynamic assessment will neither help the learner nor will it serve to compliment the information gained about learners' reading abilities from the traditional tests.

So far, the discussion concerning reading comprehension and metacognition has been largely theoretical. The next section offers a review of actual L1 and L2 dynamic reading assessment studies that have been carried out. It is followed by a review of L2 computerized dynamic reading assessment studies, with a specific focus on the structure of mediation.

Dynamic assessment of L1 and L2 reading comprehension

The idea of dynamically assessing reading ability has precedence in L1 reading literature. Cioffi and Carney (1983) first proposed the idea of using dynamic assessment to measure reading ability, specifically word-recognition, citing the work of Feuerstein, Rand, & Hoffman (1979), who used dynamic assessments with children perceived as having cognitive deficits, and Labov (1972), who developed an interesting experiment showing how a young African-American boy diagnosed as having a reading disability through a standardized assessment showed significantly different abilities in mediated assessments given in a different contexts. In reference to the Labov study, Cioffi and Carney (1983) state:

> [The child's] performance on the dynamic administration demonstrates more substantial ability consistent with his present grade placement than does the standard administration, and reveals the kinds of instruction he should receive in order to maintain that performance. A standard administration of the DRS (and most other standardized instruments) limits examiner and student interaction, often masking this important information. (p. 768)

Since the 1990s, a variety of dynamic assessment experiments related to L1 reading have appeared in the literature, though most have been related to lower-level processing skills like word recognition or phonetic decoding. For example, Carney and Cioffi (1990) discuss a word recognition dynamic assessment of one learner, Gary, a below-reading level 11-year old boy in the 5th grade who was offered assistance on a graduated scale while he took a standardized, leveled word recognition test. With instructional support, Gary was capable at reading at a much higher level — at his grade level — than standardized tests suggested. Spector (1992) conducted a

dynamic assessment of phonetic awareness with non-reading kindergarten-aged children. The purpose of the assessment was to determine whether a dynamic assessment of phonetic awareness could be a better predictor of future reading skill than other static measures of reading ability. Results of multiple regression analyses showed the dynamic assessment to be "a better predictor of kindergarten reading progress than any of the three static measures of phonemic awareness" (Spector, 1992, p. 359).

Dynamic assessment studies of more comprehensive reading measures, such as reading comprehension, are few, and most are recent and related to L2 reading, not L1. A good reason for this focus might be that L1 reading is often studied in young and beginner learners for whom reading comprehension tests are not feasible or possible. Likewise, at least in L2 English, the opposite is true — many learners are older and often tested on reading comprehension through standardized and classroom tests.

Kozulin and Garb (2002) were the first to use dynamic assessment for L2 reading comprehension. They used dynamic assessment to estimate the reading comprehension potential of at-risk high-school EFL learners in Israel. They found that dynamic assessment provided a more nuanced view of learners' abilities than did standard tests. Ajideh, Farrokhi, and Nourdad (2012) conducted a dynamic assessment of reading comprehension for a small group of university Iranian EFL learners and found that dynamic assessment provided a more detailed understanding of learners' abilities and problems than did a static assessment alone. Nourdad and Ajideh (2012) conducted a larger and more in-depth dynamic assessment of reading comprehension with university Iranian EFL learners. Nourdad and Ajideh's study included 197 participants, half of whom were part of a control group. The study included pretests, posttests, and delayed posttests for both groups. For the dynamic assessment group, five dynamic assessment sessions were held between the pre and posttest. During these sessions, learners had three standardized hints written in Persian (i.e., their L1) available for each of the reading comprehension questions. The author's found significant improvement in the reading comprehension of the dynamic assessment group over that of the non-dynamic assessment group on the posttest, reporting an effect size of $\eta^2 = .21$. There was no significant effect on the delayed posttest. Teo (2011) carried out a dynamic assessment of reading comprehension with a group of five Taiwanese university EFL learners. She found that the dynamic assessment helped distinguish learner's reading proficiency and also led learners to become more proficient, especially in terms of making inferences from reading.

Computerized dynamic assessment of reading comprehension

Aside from traditional dynamic assessments of reading comprehension, there have also been three computerized dynamic assessments of reading comprehension (Poehner & Lantolf, 2013; Shabani, 2012; Teo, 2012). Shabani's (2012) was a small-scale assessment with one reading text, one question, and ten response options, for 100 EFL learners at an Iranian university. Drawings

related to the texts' meaning were used as mediational devices. Shabani's main finding was that, without mediation provided by the computer, learners' scores would have shown a floor effect. This would have masked the differences between their abilities, which appeared more clearly through the mediation provided in the dynamic assessment. Teo's (2012) 10-week study involved 68 EFL learners at a Taiwanese university. Learners took a pretest and posttest to examine growth. For the computerized dynamic assessment participants worked on reading comprehension texts taken from TOEFL exam samples for a period of eight weeks in between the pre and posttests. The test was mediated via four hints for each question, specifically focused on mediating inferential reading skills (explained in the next section below). Teo found significant differences between pretest and posttest scores. Poehner and Lantolf (2013) conducted computerized dynamic assessments of reading and listening comprehension with learners of Chinese and French at an American university. Unlike Shabani (2012) and Teo (2012), the Poehner and Lantolf employed a *cake* format (Sternberg & Grigorenko, 2002), which means that, instead of having a pretest and posttest with mediation in between (referred to as *sandwich* format by Sternberg & Grigorenko, 2002), the mediation is integrated within the actual test. In other words, all items have mediation — there is no unmediated pretest and posttest. To determine findings, Poehner and Lantolf (2013) calculated a Learning Potential Score (LPS) (Kozulin & Garb, 2002) which discriminated between high, mid, and low level responses to the mediation. The authors also incorporated transfer texts into the assessment, thus checking learners' ability to respond to mediation on more difficult texts. The authors found that LPS "has promise as a predictor of learning" (Poehner & Lantolf, 2013, p. 336) based on statistical results comparing scores on transfer items. More detailed information about Poehner and Lantolf's study is offered in the section that follows.

Taken together, these three computerized dynamic assessment studies can provide a useful starting point for designing a computerized dynamic assessment. For example, as noted earlier, one design concern is what types of texts to use. One logical choice for dynamic assessment proponents is to use items that resemble those appearing on standardized English language learner proficiency tests; this is exactly what has taken place so far in computerized and non-computerized dynamic assessments of reading comprehension. Teo (2011) used texts and test items from TOEFL exam samples in a dynamic assessment of university level Taiwanese EFL learners, as did Teo (2012) in a computerized dynamic assessment of reading comprehension for university Taiwanese EFL learners. Poehner and Lantolf's (2013) listening and reading comprehension computerized dynamic assessment for university French and Chinese foreign language learners at an American university used texts followed by multiple choice questions, a typical standardized test format, though the exact source of the texts and items was unspecified. Poehner and Lantolf did note that all test questions and answers were in participants' L1 English "to avoid a situation wherein learners comprehend the passage but encounter difficulties with the language in which the questions and responses are framed" (p. 330).

These studies thus offer ideas both on the types of texts that might be used as well as the importance of the text item design. Concerning which computer software is useful for designing computerized dynamic assessments, Teo (2012, p. 13) reports using a "user-friendly Viewlet Quiz 3 software", while Poehner and Lantolf's (2013) test was designed by personnel at their university. Shabani (2012) gives no information about what system was used.

Computerized dynamic assessment of reading comprehension: Effective mediation

The idea that dynamic assessment can reveal more about learners' reading proficiency while simultaneously improving that proficiency is very exciting. Nevertheless, there are two major problems for researchers, at least if dynamic assessment is going to be used more widely as a complimentary approach to standard static tests. First, psychometric concerns must be addressed (Embertson, 2000). It is possible that certain growth models combined with computerized assessment may satisfy such concerns (Embertson, 2000). A second issue, mentioned earlier, lies in the mediation offered to learners during the assessment. Given that a computerized dynamic assessment must be composed of standardized hints, or graduated prompts (Brown & Ferrara, 1985), the content of mediational prompts will be quite influential on the outcome of the dynamic assessment. This section reviews the types of prompts used in the three computerized dynamic assessments cited above. What did the mediation look like? Each study is looked at in turn.

Shabani (2012) offered little rationale and specific information about the prompts used. There were four levels of prompts given when learners answered a question incorrectly. Prompts were ordered from implicit to explicit, as is normal in dynamic assessment. The four prompts in Shabini's computerized dynamic assessment are as follows:

1. Read the text again.
2. Look at the first picture and read the text again.
3. Look at the second picture and read the text again.
4. Look at the third picture and read the text again. (p. 325)

The reading text used in the assessment (Shabani, 2012, p. 328) was about a car and motorcycle accident. The pictures showed a simple depiction of the accident. The first was a drawing of the situation before the accident (i.e., two cars and a motorcycle traveling on a road). The second was the same drawing, right before the accident. The final picture was of the actual accident. The question for the text was to choose one sentence that summarized the text from among ten choices. The rationale for using pictures, or for using ten choices, or for the selection of the text was not given. However, in reference to the pictures, the author wrote that "it was assumed that pictures normally provide more information than the text" (p. 325), but no discussion was offered about why this might be true.

Teo (2012) elaborates more on the mediation used in her computerized dynamic assessment. Like Shabani (2012), she includes four levels of prompts from implicit to explicit. However, significant detail is given about the prompts. Core excerpts from Teo's prompts are as follows:

1. ...definitions of keywords in the passage as well as some common places where main ideas can be found are provided to help the learners identify the main idea...
2. ...hints are narrowed down to guide the learners to focus on certain paragraphs, or sentences, while looking for the correct answer. This more specific information is followed by an explanation of the overall meaning of the specific context...
3. ...the mediation focuses on ONE sentence, phrase, or word. The explanation given at this level is very context-specific, instead of emphasizing the overall meaning of the entire passage or specific paragraphs/sentences. The goal is to pinpoint for the learners how the specific information can lead to the correct answer...
4. ...At this level the correct answer is provided and then is followed by a step-by-step explanation... (p. 14)

Teo (2012) also gives examples of each type of mediational move described above. Thus, from a detail standpoint, the mediational interventions in Teo's study are clear. However, the underlying rationale or theory behind these specific moves is not described. In contrast, to Shabani's (2012) study with one text and ten response possibilities, in Teo's assessment, participants read three texts each week for eight weeks, with each test having one inferential reading question, and each question followed by five response possibilities.

In Poehner and Lantolf's (2013) computerized dynamic assessment of listening and reading comprehension for Chinese and French, pilot testing was done to determine proper mediation for reading test items. The authors write that "this process allowed for the identification of prompts and mediating moves (e.g., narrowing the search space within a text) that proved helpful and that were subsequently written into the C-DA program" (p. 330-331). It is unclear how many participants were involved in the pilot testing phase. After piloting, the authors designed prompts which served to "focus learner attention on the key portion of text where the correct response is found" (p. 331). This coincides well with what the mediation moves accomplished in Teo's (2012) study. However, Poehner and Lantolf (2013) offer more description for their mediational choices. They remark that "to the extent possible, the prompts are sensitive to the particular construct that an item is intended to assess. For reading comprehension the constructs are lexis, grammar, discourse, and culture" (p. 330). Their study also differs from others in putting all the questions, responses, and mediational prompts in learners' L1 (i.e., English). As with Teo (2012), the texts used in Poehner and Lantolf (2013) had five response options, which the authors say allowed them to "maximize the number of prompts" (p. 330). Poehner and Lantolf (2013) describe the prompts

they use as follows:

> Throughout the tests, an initial incorrect response is met with the following statement, presented aurally in the listening test and in writing in the reading test: 'Sorry. That's not the right answer.' The first prompt is then presented, either by highlighting in colored font or by auditory repetition of a chunk of text that the learner should pay attention to. The learner then selects from among the remaining four response options. Following a second incorrect response the learner is again informed that the response is incorrect: 'Sorry. That's not right either.' At this point the scope of the search space is further narrowed. This process continues for a total of four attempts, at which point the correct answer is revealed accompanied by an explanation. At any point where the correct response is selected the option of viewing an explanation supporting the correct response is provided. This procedure is intended to help learners who answer correctly but who may be uncertain of their response or who guess the correct answer. (p. 331)

Poehner and Lantolf's (2013) system of prompts shows more depth than those discussed in Shabani (2012) and Teo (2012). They also go further than both of those studies in integrating another important tenet of dynamic assessment – transfer. Poehner and Lantolf define transfer as "introducing assessment tasks that follow the same principles as earlier ones but are more difficult or complex [to] offer insights into whether learners have internalized mediation previously offered" (p. 326). In their study, there were five transfer items on both the French assessments and on the Chinese listening assessments. There were eight transfer items on the Chinese reading test. Transfer texts were not identified as such for learners. They only differed in terms of difficulty. While the authors offer a detailed explanation of how transfer texts differed from non-transfer texts, it will not be rewritten here. It is sufficient to say that transfer texts were more difficult in terms of content and/or lexical density, and that the purpose of the transfer texts was "to ascertain whether learners show signs of improvement in recognizing and interpreting words and phrases in the target language more generally (an important component of reading ability) over the course of the test" (p. 333).

This summary of three computerized dynamic assessments' mediational moves is useful for understanding where to begin in designing a computerized dynamic assessment of reading comprehension. Mediation in Teo (2012) and Poehner and Lantolf's (2013) dynamic assessments share general features, but the latter authors provide a more in-depth approach to the assessment. Considering all of the studies, a good computerized assessment of reading comprehension will probably include a number of things, as follows:

1) Numerous mediational moves, perhaps gradually narrowing the area that readers must

attend to in order to answer questions. Highlighting might be employed to help guide learners to those areas.

2) At least five choices if using multiple choice responses.
3) A number of transfer items of increasing difficulty in comparison to non-transfer items.

Concerns about computerized dynamic assessment

Regardless of how the three guidelines above are followed, piloting is an important procedure. Poehner and Lantolf's (2013) mention piloting, but the pilot study is not described. Piloting is an important process in the development of a strong testing instrument. In all the computerized dynamic assessment studies reviewed, detailed discussion of the test instruments is missing. What makes good reading comprehension items? What makes good distractors? There is ample research in this area, and it is a missing component of the L2 dynamic assessment studies to date. Teo (2012) did mention using TOEFL sample readings because they were professionally designed and validated, but some text analysis and validation among learners similar to those being assessed would further support and contextualize their usage. The validity and reliability of assessment items is essential since a computerized dynamic assessment may only be as good as the instruments used in it. Using standardized test samples might be a useful strategy, but it remains important for assessors to offer specific profiles of the items that are used. Naturally, it might sometimes make sense for researchers to develop their own items. In this case as well, a clear rationale for the design of the items will be important for L2 computerized dynamic assessment research to advance.

Another important concern that should be addressed is the effects of guessing and the increased possibility that learners can get an answer correct as response options decrease. In other words, if a multiple choice question first has five responses, and then the learner answers incorrectly and receives mediation, then they will have an easier task of choosing from just four options. Poehner and Lantolf (2013) accede that this may be an issue:

> Allowing learners multiple attempts at each test item decreases the degrees of freedom they confront and therefore increases the possibility of guessing the correct answer. The reduction in the degrees of freedom after failed attempts means subsequent mediation might be doing less 'work' with regard to leading learners through a comprehensive process. (p. 338)

The authors go on to remark that there still can be significant benefit from the assessment, and that multiple choice questioning inherently might involve guessing. Still, the issue is something that an assessment designer should consider. Shabani's (2012) strategy of having ten response options does not seem practical, but considering other ways to test comprehension could be useful.

Conclusion

This paper's purpose has been to review computerized dynamic assessment of reading comprehension and postulate procedures for future design of such assessments. Recent attempts at computerized dynamic assessments of reading comprehension offer a rough guide for those wishing to develop new assessments. However, they still leave many unanswered questions, such as best practices for mediation, and appropriate content for items. Computerized dynamic assessment is still an emerging approach to assessment, but if valid procedures for the creation and administration of computerized dynamic assessments of reading comprehension and other skills are realized, it can add useful perspective to information gained through other standardized, static tests.

References

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher, 61* (5), 364-373.

Ahmadi, M. R., Ismail, H. N., & Abdullah, M. K. K. (2013). The importance of metacognitive reading strategy awareness in reading comprehension. *English Language Teaching, 6* (10), 235-244.

Ajideh, P., Farrokhi, F., & Nourdad, N. (2012). Dynamic assessment of EFL reading: Revealing hidden aspects at different proficiency levels. *World Journal of Education, 2* (4), 102-111.

Alderson, J. C. (2000). *Assessing reading.* Cambridge: Cambridge University Press.

Blachowicz, C. (1999). Vocabulary in dynamic reading assessment: Two case studies. *Journal of Reading Psychology, 20,* 213-236.

Bridges, M. S., & Catts, H. W. (2011). The use of dynamic screening of phonological awareness to predict risk for reading disabilities in kindergarten children. *Journal of Learning Disabilities, 44,* 330-338.

Brown, A., & Ferrara, R. A. (1985). Diagnosing zones of proximal development. In J. V. Wertsch (Ed.), *Culture, communication and cognition: Vygotskian perspectives* (pp. 273-305). Cambridge: Cambridge University Press.

Carney, J., & Cioffi, G. (1990). Extending traditional diagnosis. The dynamic assessment of reading abilities. *Reading Psychology, 11* (3), 177-192.

Carney, J., & Cioffi, G. (1992). The dynamic assessment of reading abilities. *International Journal of Disabilitiy, Development, and Education. 39* (2), 107-114.

Cioffi, G., & Carney, J. (1983). Dynamic assessment of reading disabilities. *The Reading Teacher, 36,* 764-768.

Davin, J. K. (2013). Integration of dynamic assessment and instructional conversations to promote development and improve assessment in the language classroom. *Language Teaching Research, 17,* 303-322.

Dörfler, T., Golke, S., & Artelt, C. (2009). Dynamic assessment and its potential for the assessment of reading competence. *Studies in Educational Evaluation, 35,* 77-82.

Feuerstein, R., Rand, Y., & Hoffman, M. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques.* Baltimore, MD: University Park Press.

Fuchs, D., Compton, D., Fuchs, L., Bouton, B., and Caffrey, E. (2011). The construct and predictive validity of a dynamic assessment of young children learning to read: Implications for RTI frameworks. *Journal of*

Learning Disabilities, 44, 339-347.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice.* New York: Cambridge University Press.

Guterman, E. (2002). Toward dynamic assessment of reading: Applying metacognitive awareness guidance to reading assessment tasks. *Journal of Research in Reading, 25*, 283-298.

Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice.* New York: Cambridge University Press.

Heatley, A., Nation, I.S.P. and Coxhead, A. (2002). RANGE and FREQUENCY programs [computer software]. Retrieved from: http://www.vuw.ac.nz/lals/staff/Paul_Nation

Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL test comprehension. *School Psychology International, 23*, 112-127.

Kozulin, A., & Garb, E. (2004). Dynamic assessment of literacy: English as a third language. *European Journal of Psychology of Education, 19* (1), 65-77.

Labov, W. (1972). *Language in the inner city.* Philadelphia, PA: University of Philadelphia Press.

Lantolf, J. P., & Poehner, M. E. (1994). Dynamic assessment: Bringing the past into the future. *Journal of Applied Linguistics, 1*, 49-74.

Lantolf, J. P., & Poehner, M. E. (2008). In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of Language and Education, 2nd Edition, Volume 7: Language Testing and Assessment* (pp. 273-284). Springer: New York.

Larson, J. A., & Nippold, M. N. (2007). Morphological analysis in school-age children: Dynamic assessment of a word learning strategy. *Language, Speech, and Hearing Services in Schools, 38* (3), 201-212.

Michalsky, T., Mevarech, Z. R., & Haibi, L. (2009). Elementary school children reading scientific texts: Effects of metacognitive instruction. *The Journal of Educational Research, 102* (5), 363-374.

National Reading Panel (NRP) (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.* Bethesda, MD: NICHD.

Nourdad, N., & Ajideh, P. (2012). The immediate and delayed effect of dynamic assessment of EFL reading ability. *English Language Teaching, 5* (12), 141-151.

Paris, S. G. (2007). Assessment of Reading Comprehension. *Encyclopedia of Language and Literacy Development* (pp. 1-8). London, ON: Canadian Language and Literacy Research Network. Retrieved from http://www.literacyencyclopedia.ca/pdfs/topic.php?topId=226

Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development.* Berlin: Springer.

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.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95*, 26-43.

Schneider, E., & Ganschow, L. (2000). Dynamic assessment and instructional strategies for learners who struggle to learn a foreign language. *Dyslexia, 6*, 72-82.

Shabani, K. (2012). Dynamic assessment of L2 learners' reading comprehension processes: A Vygotskian perspective. *Procedia – Social and Behavioral Sciences, 32*, 321-328.

Spector, J. E. (1992). Predicting progress in beginning reading: Dynamic assessment of phonemic awareness. *Journal of Educational Psychology, 84*, 353-363.

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

Teo, A. (2011, Sep.). Using dynamic assessment to promote college EFL learners' reading skills in Taiwan. Paper presented at the Educational Research Association in Singapore (ERAS), Singapore.

Teo, A. (2012). Promoting EFL students' inferential reading skills through computerized dynamic assessment. *Language Learning & Technology, 16* (3), 10-20.

(Received February 18, 2016)