# Computer-based Assessment Vs. PaperPencil-based Assessment: An Investigation into the Performance and Attitude of Iranian EFL Learners' Reading COmprehension

**Article** · January 2014

CITATIONS
11

READS
2,126

**4 authors**, including:

Hamid Mojarrad
Payame Noor University
**2** PUBLICATIONS **16** CITATIONS
SEE PROFILE

Fatemeh Hemmati
Payame Noor University
**24** PUBLICATIONS **81** CITATIONS
SEE PROFILE

Amir Sadeghi
University of Canterbury
**7** PUBLICATIONS **23** CITATIONS
SEE PROFILE

International Journal of Language Learning and Applied Linguistics World
(IJLLALW)
Volume 4 (4), December 2013; 418-428 Mojarrad, H., et al
ISSN (online): 2289-2737 & ISSN (print): 2289-3245 www.ijllalw.org

# COMPUTER-BASED ASSESSMENT (CBA) VS. PAPER/PENCIL-BASED ASSESSMENT (PPBA): AN INVESTIGATION INTO THE PERFORMANCE AND ATTITUDE OF IRANIAN EFL LEARNERS' READING COMPREHENSION

**Hamid Mojarrad**
*Department of linguistics and foreign languages, Payame Noor University, POB: 19295-4697, Tehran, I.R. Iran*
*E-mail: hamidmojarrad@gmail.com*

**Fatemeh Hemmati**
*Payame Noor University, Iran.*
*E-mail: hematitefl@gmail.com*

**Manuchehr Jafari Gohar**
*Payame Noor University, Iran.*
*E-mail:jafarigohar2007@yahoo.com*

**Amir Sadeghi**
*University of Canterbury, New Zealand.*
*E-mail: amir.sadeghi@canterbury.ac.nz*

**ABSTRACT**
To examine the impact of transitioning traditional reading comprehension assessments to computer-based, 66 male English as a Foreign Language (EFL) learners aged 8 to 12 years were assigned to take two different twenty-minutes reading comprehension tests with the same level of difficulties on paper and computer screen using scrolling text to navigate through pages. They also completed an attitude questionnaire to reveal their attitudes towards computerized testing. The findings revealed that there are no significant differences in reading comprehension scores across testing modes. In addition, evaluating the Likert type questionnaire revealed that the majority of students prefer to take the test on computer. The findings also suggest that the amount of reading comprehension among children does not differ considerably while switching from paper/pencil-based assessment into computer-based assessment. Hence, schools with no or limited computers and the Internet facilities should not be concerned about the students' performance and their level of achievement.

**KEYWORDS:** computer-based assessment, paper/pencil-based assessment, attitude.

## INTRODUCTION

Computerized testing began in the early 1970s (Drasgow, 2002; Wainer, 1990). Limited computer capability and high costs, however, used to limit the implementation of computerized testing. With the advent of new technologies, computerized testing has begun to be developed and implemented in large-scale testing programs such as licensure, certification, admissions, and psychological tests (Kim & Huynh, 2007). For example, the Graduate Record Examinations (GRE) has been administered in computer-adaptive format for several years. Likewise, in 1998, the Test of English as a Foreign Language (TOEFL) began transitioning to computer-adaptive testing. Recently, the new TOEFL Internet-Based Test (IBT) began administration via the Internet in a non-adaptive format. Increased testing requirements and tight deadlines imposed by the No Child Left Behind Act of 2001 (NCLB) (Public Law No: 107–110) have led to new ways that states can measure student performance more efficiently (Kim, et al., 2007).

The advantages of computers are well known and apparent (Al-Amri, 2009). Computers offer test developers the opportunity to improve their productivity and lead to innovation in their fields. The standardization of test administration conditions is one of the benefits offered by Computer-Based Testing (CBT). No matter what the tests' population size is, CBT helps test developers to set the same test conditions for all participants. Al-Amri (2009) also believed that CBT improves all aspects of test security by storing questions and responses in encrypted databases and enables testers to create randomized questions and answers from vast question pools. Moreover, offering different test formats and the immediate presentation of different types of feedback, either to students or to testers, are other great advantages of CBT. Collecting different performance data such as latency information is a unique feature of CBT (Olsen, Maynes, Slawson, & Ho, 1989). On the examinees' side, they are able to receive greater measurement efficiency and the possibility to take the test at any time. On the other hand, there are some disadvantages that users have to be aware of before opting for computer-based testing, which led many scholars to suggest conducting systematic studies to check equivalency and comparability of paper-based tests and computer-based tests (Parshall, Spray, Kalohn, & Davey, 2002). For example, students need some degree of computer literacy in order to avoid the mode effect on computer-based testing (Alderson, 2000).

## LITERATURE REVIEW

As there has been a growing interest in computer-based testing in large-scale assessments, several comparability studies have involved elementary and secondary students over the past few years. Russell and Haney (1997) investigated the mode effects on middle school students' performances on open-ended items in writing, science, math, and reading, as well as multiple-choice and short-answer items in language arts, science, and math from the National Assessment of Educational Progress (NAEP), and extended writing test items. They found that the effect of administration mode was not significant for the multiple-choice items, but a substantial effect was found for the open-ended items. The results showed that students who were used to writing by the computer performed better when they responded to the open-ended test using a computer rather than using pencil and paper.

Similar findings were found in studies by Russell (1999) and Russell and Plati (2001), who reported that students who were accustomed to writing using a computer performed better on the open-ended tests when they wrote using a computer than when they wrote by hand. Pommerich (2004) investigated the item-level mode effects of English, reading, and science reasoning tests in grades 11 and 12 and found that examinees responded differently to some items under the various interface features, although the mode effect in general was small. Pommerich found that the paging condition group outperformed the scrolling condition group in the reading and science reasoning tests, and the automatic scrolling group performed slightly better than the semi-automatic scrolling group in the English test.

Two comparability studies on the online versions of the NAEP math and writing tests showed that the paper group significantly outperformed the computer group in the eighth-grade NAEP mathematics test, but no mode effect was found for the eighth-grade NAEP essay test (Sandene, Horkay, Bennett, Allen, Braswell, Kaplan, 2005). The NAEP studies also found that students' familiarity with computers was related to their performance. Particularly, hands-on measures of keyboarding skill were found to be a significant predictor of students' performances on the NAEP online writing test (Sandene, et al., 2005). Although the NAEP studies have directly investigated administration mode effects in the K–12 large scale assessments, the NAEP is a low-stakes assessment and the lack of consequences for its results could affect student motivation to take the test seriously, and the results may not be generalized to high-stake statewide assessments.

Regarding the comparability of these two assessment modes on children as the participants, Barnes (2010) examined the feasibility of using CBA with children who have not yet started the first grade. Kindergarten children were able to perform the assessment nearly as well on the CBA as they did on the traditional paper version of the test. However, most children under four years of age were not able to complete a CBA independently. The children reacted positively to the CBA and seemed to enjoy using the computer, but the test results indicate that the CBA was more difficult and may be measuring something more than rhyme awareness for the preschoolers. In 2005, Sim examined the comparability of CBA and PPBA on 20 children, aged between 7 and 8 of mixed gender. The results showed that the majority of children performed better on paper than computer although there was no significant difference.

In fact, a lot of research works have been conducted to evaluate the comparability of computer based assessment and paper and pencil based assessment. Some studies revealed that there is a significant difference between the two testing modes on test scores (e.g. Scheuermann & Björnsson, 2009; Choi, Kim, & Boo, 2003), while other studies reported opposite or inconsistent results (e.g. Al-Amri, 2009; Boo, 1997). However, unlike the abundance of CBA research done with older or special needs students, there is a dearth of available research focusing on the issues of computer-based assessment with typically developing young children (Barnes, 2010). Therefore, a practical comparison between these two methods needs to be done in order to identify whether the testing mode has considerable influence on the performance of children when they sit for a reading comprehension assessment. In other words, due to the scarcity of empirical research concerning language assessment and use of technology in language

assessment in Iran, this study hopes to raise awareness on the use of technology in language assessment in Iranian context.

## RESEARCH QUESTIONS

Specifically, the research questions were:

1. Is there any significant difference between computer-based assessment and paper and pencil based assessment when assessing reading comprehension among primary school students?
2. What is the test takers' attitude towards computer-based assessment compared with traditional assessment?

## METHODOLOGY

### *Participants*

This study recruited 66 male EFL learners aged 8 to 12 years from one of the well-known language learning institutes to take part in PPBA and CBA as the participants. Since the numbers of students who were at the required level of proficiency and appropriate for this study were not high enough to employ random selection, all of the students who had the requisites to sit for the reading compehesion assessments were selected as the participants. They were chosen from among five different proficiency levels ranging from *High Beginning* to *Low Advanced*. The number of students per class ranged from 5 to 23. It is necessary to notify that every student sit for the required placement tests at the very beginning of his entry into the institute. Hence, there were no worries about the proficiency levels of the children.

### *Instruments*

Two reading comprehension tests as well as an attitude questionnaire were utilized in this study as the data collection instruments, which are described below in details.

*Reading Comprehension Test*

Based on the proficiency levels of the participants, two different tests with the same level of difficulty containing a reading passage, a short story, and some associated multiple choice items were extracted directly from released *Read Theory* or *English for Every One* organizations. The tests for PPBA and CBA, different in terms of content but similar in terms of difficulty level as well as number of paragraphs and question items, were also checked by the institute teachers to ensure from their appropriateness.

With regard to the CBA, some points are necessary to be referred to in advance. Firstly, the reading comprehension tests were embedded in a dynamic web page with a submit button at the bottom of the page to send the students' entered answers to the website database. Secondly, in order to avoid building a separate website from scratch, the researchers preferred to benefit from the existing online form builders such as Google Drive, Adobe Acrobat Form Creator, and so on

which not only allow to use their templates for creating the desired pages, but also offer facilities to the developer to analyze the entered data.

*The Attitude Questionnaire*

After completing the web-based reading comprehension assessment, all students were asked to complete a web-based questionnaire that consisted of five question items. The purpose of this questionnaire was to measure students' attitude toward computer-based assessments. The questions were addressing the extent to which the use of computers has considerable influence on the amount of comprehension when children sit for the reading exams. Since the participants were children of eight to 12 years and there was a great fear about their perception and understanding of the question items, translating the items into their native language (Persian) would resolve this concern and prevent from any kinds of misinterpretation.

A good number of similar questionnaires from the previous studies (e.g. The National Assessment of Educational Progress, 2007; Documents & Resources for Small Businesses & Professionals, 2010; Melhado, 2010) were evaluated to extract the appropriate items for children. Then, the selected items were merged in a well-designed format and checked by the institute's teachers as well to ensure from their appropriateness and the validity of items. The questionnaire was designed according to the Likert scale with 5 question items and four possible answers for each one including *Strongly Agree*, *Agree*, *Disagree*, and *Strongly Disagree*. In other words, based on some researchers' idea (e.g. Chang, 1994; Cronbach, 1950; Adelson & McCoach, 2010; Kulas, Stachowski, & Haynes, 2008), the *Neutral* (*Neither Agree nor Disagree*) option was omitted from the middle of the scale in order to avoid children having spurious answers and instead, specify exactly their viewpoints towards the questions.

To ensure its reliability, the Cronbach's alpha reliability coefficient was calculated. Based on Cronbach's formula, the researcher achieved an alpha value of 0.83, which indicates a very acceptable degree of reliability to be utilized in the study.

***Procedure***

The process of data collection for this study was done during May 2013. Students were given twenty minutes to complete the reading comprehension tests in two attempts of administration, one for PPBA and the other for CBA. At the end of CBA, the students were also given an additional 10 minutes to complete the attitude questionnaire. To control for effects that might result from differences in the computers available within each class, a set of MSI 11-inch laptops were brought into the classes with no external mice. The laptops were chosen among the small and light weighted sizes to be appropriate for children as the participants.

## RESULTS AND DISCUSSION

In order to answer the first and second research questions, a paired-samples t-test at 0.05 levels of significance was conducted and the frequencies of each item in the questionnaire were calculated respectively. The descriptive statistics for PPBA and CBA is depicted in Table 1.

*Table 1: Descriptive statistics for PPBA and CBA*

| Assessments | Mean | N | SD |
|---|---|---|---|
| Paper/Pencil based Assessment | 57.29 | 66 | 19.56 |
| Computer based Assessment | 63.33 | 66 | 23.70 |

*Note: N=Number of the students; SD=Standard Deviation*

Table 2 presents the output of Paired-Sample t-test.

*Table 2: Paired-Samples T-test Results of comparison between PPBA and CBA scores*

| | **Mean** | **SD** | **t** | **df** | **Sig.** |
|---|---|---|---|---|---|
| Paper – Computer | -6.03 | 26.78 | -1.832 | 65 | .072 |

*Note: SD=Standard Deviation; Sig. =Significance value; df stands for Degrees of freedom. *P < .05, **P < .01*

The results of Paired-Sample t-test suggest that the amount of reading comprehension among children does not differ considerably while switching from traditional mode of testing into the computerized one.

To analyze the questionnaire, the researcher made use of only the simple comparisons of frequency percentages related to the predefined answers for the questionnaire's questions, namely, *Strongly Agree*, *Agree*, *Disagree*, and *Strongly Disagree*. This measure identified that which item has the most frequency among the others and consequently determined the students' attitude toward CBA. Table 3 and figure 1 represent the descriptive statistics and the column chart associated with the questionnaire respectively.

*Table 3: Descriptive statistics for attitude*

| Mean | SD | N |
|---|---|---|
| 17.78 | 2.54 | 5 |

*Note: SD=Standard Deviation; N=Number of items*

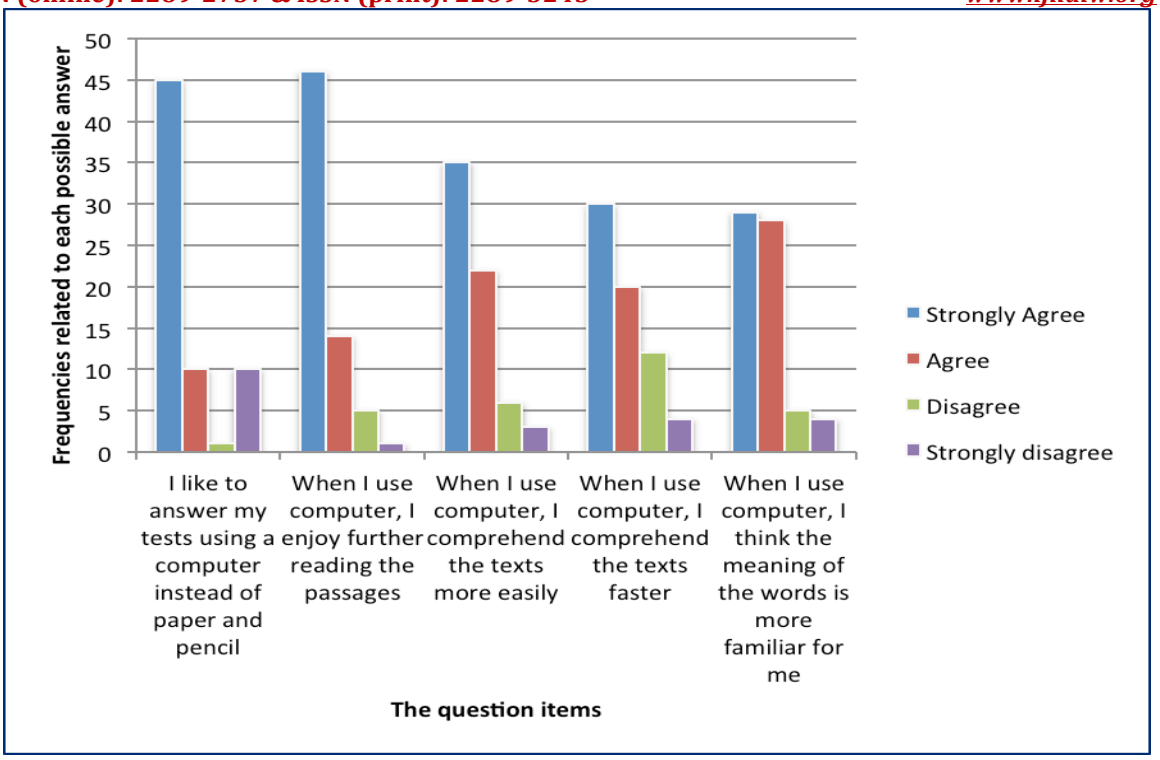


*Figure 1: Column chart showing frequencies of each possible answer of the Likert-type questionnaire*

Obviously, the first two possible answers, *Strongly Agree* and *Agree*, had the most frequencies in each question, so it is simply concluded that the participants had better feeling and attitude towards administering the assessments using computers. In addition, the researcher's observation at the time of CBA is another evidence for this claim.

Regarding the first research question, the analysis of the results proved that the kind of administration either paper-based or computer-based will not significantly influence the participants' performance in reading comprehension assessments and the existing differences are not considerable and salient but due to chance or other external factors.

This finding is in line with Higgins, Russell, and Hoffmann (2005)'s study who investigated the Effect of Computer-Based Passage Presentation on Reading Test Performance and found that there were no statistically significant differences in reading comprehension scores based on computer fluidity and computer literacy. Likewise, the study carried out by Baumer, Roded, and Gafni (2009) on the Equivalence of Internet-Based vs. Paper-and-Pencil Psychometric Tests, revealed that the modality of administration does not affect test performance. Pommerich (2004) also conducted a research on English, Reading, and Science passage-based tests and found that there were no significant differences in scores across modes for the Reading test. In addition, Al-Amri (2009) investigated the comparability of computer-based testing versus paper-based testing and found that testing mode has almost no significant effect on the overall validity and reliability of the tests. He also asserted that computer familiarity has no influence on students' performance

and the other factors such as computer attitude and prior testing mode preference do not have any critical impact on the overall students' performance.

However, this finding may be at odds with some previous studies as well. For instance, Fitzpatrick and Triscari (2005) in their study found that item parameters and cut score results are comparable across administration modes. Likewise, the study performed by Choi and Tinkler (2002) revealed that the computerized Oregon statewide reading and mathematics tests were more difficult for third graders, but the paper version of the test was more difficult for tenth graders. They also found that mode effects were more pronounced in reading tests than in mathematics tests. Similar findings were reported in the study by Way, Davis, and Fitzpatrick (2006), who investigated the comparability of paper and online versions of the Texas statewide tests in mathematics, reading/English language arts, science and social studies at grades 8 and 11. The results showed that the tests were more difficult for the online group than for the paper group and the administration mode effects were more evident for ELA than for other subjects.

Regarding the second research question, the analyses of the attitude questionnaire as well as the researcher's observation at the time of CBA administration identified that children had a better sense and attitude toward taking the examinations using computers.

This finding is consistent with the results of many previous related studies. For instance, Higgins, Russell, and Hoffmann (2005) in their study, which was referred to earlier, examined the attitude of the participants as well. They asked the participants whether they thought it was easier or harder to take the test on computer or paper. Of the 135 students who responded to this item, 82% reported that it was easier to take the test on computer. In addition, students were asked in a selected response format if they would have preferred to take the test on computer or on paper. Of the 161 students who responded to this item, 87% reported that they would prefer to take the test on computer. Likewise, Way, Davis & Fitzpatrick (2006) believed that according to recent surveys, students tested online feel comfortable with taking tests on the computer and tend to prefer it to traditional paper testing. Moreover, the research work done by Chua (2012) showed that the computer-based testing has effectively increased intrinsic and extrinsic motivation of the test takers in challenge, curiosity, self-efficacy, involvement, joyfulness, comprehension and social dimensions. However, answering the test in a shorter time with higher testing motivation level did not help a test taker to achieve a higher score.

The majority of previous studies are in agreement with the above finding. However, there are some other studies which are in contrast with the current study. For example, the results of the study done by Ward, Hooper, & Hannafin (1989) indicated a negative attitude toward computer testing so that seventy-five percent of the computer-tested group either strongly agreed or agreed that computer testing was more difficult than traditional methods. In addition, Durndell & Thompson (1997) and Whitely (1997) believed that females usually have more negative attitudes toward computers than males. Moreover, Sam, Othman & Nordin (2005) in their study found that the respondents had moderate computer anxiousness and medium attitudes toward the Internet. Yet, it is not extravagant to say that almost all of the recent studies have reached the same conclusion as that of the current study; because with the vast development of technology and its

usage in the lives of people around the world, the students' positive attitudes towards computerized assessment is not unexpected.

## CONCLUSION

Based on the above-mentioned discussion, it can be concluded that although the mode of testing does not noticeably influence the amount of children's reading comprehension, the teachers, schools and institutes can benefit from technology particularly the computers and internet in their educational environments to promote and encourage students towards assessments. Furthermore, those schools and institutes, which do not have required facilities to bring technology in the classrooms, should not be worried about the students' performance and their level of achievement.

### *Limitations and Implications*

Due to the nature of the present research, which necessitated providing similar PCs or laptops for a number of students in order to carry out computer-based assessment, there were some limitations in some aspects of this study such as the small sample size, lack of female participation, low speed of internet, and so on. Hence, future research should be conducted on a larger and more diverse sample of students and should be expanded to include students in other grade levels. The number of passages and items should also be increased. The future study with a larger and more diverse sample and more items may show more subtle differences in the performance of participants particularly when female students are also taken part in the study.

The current research can be incorporated into regular classroom settings and English language learning institutes with the intention of improving the quality of learning, teaching, and testing. In the case of classroom scale assessment, teachers can make informed decisions of how they can measure children's reading comprehension especially using computers and technology, which have come to the lives of children around the world in the recent decades, as the alternative tools for traditional assessments. In the case of large-scale assessment, the study may provide a basis for test designers/developers especially in measurement organizations and English language institutes to act according to the results.

## REFERENCES

Adelson, J. L., & McCoach, D. B. (2010*)*. Measuring the mathematical attitudes of elementary students: The effects of a 4-point or 5-point Likert-type scale. *Educational and Psychological Measurement*, *70 (5),* 796-807.

Alderson, J. C. (2000). Technology in Testing: the Present and the Future. *System*, *28(4),* 593-603.

Al-amri, S., (2009, February).*Computer-based testing vs. paper-based testing: establishing the comparability of reading tests through the evolution of a new comparability model in a*

*Saudi EFL context.* Thesis submitted for the degree of Doctor of Philosophy in Linguistics.University of Essex.

Barnes, S. K. (2010). *Using computer-based testing with young children.* (Ph.D. dissertation). Retrieved from ProQuest Dissertations and Theses database. (Publication No. AAT 3403029.)

Baumer, M., Roded, K., & Gafni, N. (2009). Assessing the equivalence of Internet-based vs. paper-and-pencil psychometric tests.In D. J. Weiss (Ed.), *Proceedings of the 2009 GMAC Conference on Computerized Adaptive Testing*.

Boo, J. (1997).*Computerized versus paper-and-pencil assessment of educational development: score comparability and examinee preferences.*Unpublished dissertation, University of Iowa.

Chang, L. (1994). A psychometric evaluation of 4-point and 6-point Likert-type scales in relation to reliability and validity.*Applied Psychological Measurement,* 18(3), 205-215.

Choi, I-C., Kim, K. S., & Boo, J. (2003). Comparability of a paper-based language test and a computer-based language test.*Language Testing,* 20, 295-320.

Choi, S.W., & Tinkler, T. (2002). *Evaluating comparability of paper and computer-based assessment in a K–12 setting*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, LA.

Chua, Y. (2012). Effects of computer-based testing on test performance and testing motivation. *Journal of Computers in Human Behavior* (pp. 1580-1586).

Cronbach, L. J. (1950). Further evidence on response sets and test design. *Educational and Psychological Measurement*, *10(1),* 3-31.

Drasgow, F (2002). The work ahead: A psychometric infrastructure for computerized adaptive tests. In C.N. Mills, M.T. Potenza, J.J. Fremer, & W.C. Ward (Eds.), *Computer-based testing: Building the foundation for future assessments* (pp. 67–88). Hillsdale, NJ: Lawrence Erlbaum.

Durndell, A., & Thompson, K. (1997). Gender and computing: A decade of change? *Computers and Education*, *28 (1),* 1-9.

Fitzpatrick, S., &Triscari, R. (2005, April).*Comparability studies of the Virginia computer-delivered tests*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Higgins, J., Russell, M., & Hoffmann, T. (2005). Examining the Effect of Computer-Based Passage Presentation on Reading Test Performance.*Journal of Technology, Learning, and Assessment*.

Kim, D.-H., & Huynh, H. (2007). Comparability of Computer and Paper-and-Pencil Versions of Algebra and Biology Assessments.*Journal of Technology, Learning, and Assessment*, 6(4). Retrieved April 11, 2013, from http://www.jtla.org.

Kulas, J. T., Stachowski, A. A., & Haynes, B. A. (2008). Middle response functioning in Likert-responses to personality items.*Journal of Business and Psychology*, 22(3), 251-259.

Melhado, G., (2010).*Improving Reading Comprehension Through the Use of Computer-Aided Instruction Among Third Graders.* Thesis submitted for the degree of Doctor of Education. Nova Southeastern University.

Nichols, P., & Kirkpatrick, R. (2005, April). *Comparability of the computer administered tests with existing paper-and-pencil tests in reading and mathematics tests*. Paper presented at

the Annual Meeting of the American Educational Research Association, Montreal, Canada.

Olsen, J., Maynes, D., Slawson, D., & Ho, K. (1989) Comparison of paper-administered, computer-administered and computerized achievement tests. *Journal of Educational Computing Research*, *5*, 311-326.

Parshall, C., Spray, J., Kalohn, J. & Davey, T. (2002). *Practical Considerations in Computer-Based Testing*. : New York: Springer.

Poggio, J., Glasnapp, D.R., Yang, X., &Poggio, A.J. (2005). A comparative evaluation of score results from computerized and paper and pencil mathematics testing in a large scale state assessment program.*Journal of Technology, Learning, and Assessment*, *3(6)*.

Pommerich, M. (2004). Developing computerized versions of paperand- pencil tests: Mode effects for passage-based tests. *Journal of Technology, Learning, and Assessment*, *2(6)*.

Russell, M. (1999). Testing on computers: A follow-up study comparing performance on computer and on paper.*Education Policy Analysis Archives*, *7*.

Russell, M., & Haney, W. (1997). Testing writing on computers: An experiment comparing student performance on tests conducted via computer and via paper-and-pencil. *Education Policy Analysis Archives*, *5*.

Russell, M., & Plati, T. (2001). Effects of computer versus paper administration of a state-mandated writing assessment.*TC Record.Org*. Retrieved April 19, 2002, from http://www.tcrecord.org/Content.asp?ContentID=10709

Sam, H. K., Othman, A. E. A., & Nordin, Z. S. (2005). Computer Self-Efficacy, Computer Anxiety, and Attitudes toward the Internet: A Study among Undergraduates in Unimas. *Educational Technology & Society*, *8 (4),* 205-219.

Sandene, B., Horkay, N., Bennett, R., Allen, N., Braswell, J., Kaplan, B., & Oranje, A. (2005). *Online assessment in mathematics and writing: Reports from the NAEP technology-based assessment project*, Research and Development Series (NCES 2005-457). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Scheuermann, F., & Bjo¨rnsson, J. (2009).*The transition to computer-based assessment*. Luxembourg: Officefor Official Publications of the European Communities.

Sim, G. & Horton, M. (2005). Performance and Attitude of Children in Computer Based Versus Paper Based Testing. In P. Kommers & G. Richards (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2005* (pp. 3610-3614). Chesapeake, VA: AACE.

Wainer, H. (1990).*Computerized adaptive testing: a primer*. Hillsdale, N J: LEA.

Ward, T., Hooper, S., & Hannafin, K. (1989).The Effect of Computerized Tests on the Performance and Attitudes of College Students.*Journal of Educational Computing Research* (pp. 327 - 333).

Way, W. D., Davis, L. L., & Fitzpatrick, S. (2006, April). *Score comparability of online and paper administrations of the Texas Assessment of Knowledge and Skills*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco, CA.

Whitely, B. (1997). Gender differences in computer related attitudes and behavior: A meta analysis. *Computers in Human Behavior*, *13 (1),* 1-22.