STUDENTS' PERFORMANCE AND PREFERENCE ON
COMPUTER-BASED TESTS VS. PAPER-BASED TESTS:
A CASE STUDY

by

Melinda S. Cole

A Master's Project
Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of Science in Education
Literacy Education (Birth-Grade 6)
Department of Language, Learning, and Leadership
State University of New York at Fredonia
Fredonia, New York

May 2013

State University of New York at Fredonia
Department of Language, Learning, and Leadership

CERTIFICATION OF PROJECT WORK

We, the undersigned, certify that this project entitled *Students' Performance and Preference on Computer-Based Tests vs. Paper-Based Tests: A Case Study* by Melinda S. Cole, Candidate for the Degree of Master of Science in Education, Literacy: Birth to Grade 6, is acceptable in form and content and demonstrates a satisfactory knowledge of the field covered by this project.

_____   *May 17, 2013*
Master's Project Advisor   Date
Dr. Xiaoning Chen
Department of Language, Learning, and Leadership

_____   6-5-2013
Dr. Anna Thibodeau   Date
Department of Language, Learning, and Leadership

_____   6-24-13
Dean Christine Givner   Date
College of Education
At SUNY Fredonia

Abstract

The purpose of this action research project was to determine if students performed better on computer-based tests or paper-based tests, as well as to determine the students' preference for each testing method. The participants included seven fourth grade students. Data sources included: Results from two computer-based comprehension tests and two paper-based comprehension tests, two student surveys which included five questions regarding the testing experience and opinions of the assessments, open-ended interviews with the students in small groups, and the researcher's observational notes. Data were analyzed for patterns and themes. The findings demonstrated that the majority of students performed better on the paper-based test, and that the majority of students preferred the paper-based test. Factors which made the students' performances on computer versus paper tests more complicated included: Gender, number of tests administered, socioeconomic status, and the sample size of the participants.

Table of Contents

Chapter One Introduction

Statement of the Issue

The New York State Education Department proposes that all students in grades 3-12 within the public school system will take all State and Regents examinations online, eliminating the paper and pencil examinations. In his May 2012 memorandum, New York State's Commissioner of Education, John B. King, discussed the reasons for switching to computer-based testing for public schools in New York State. King emphasized the need for computer-based testing due to our technology-driven world. According to King, in order for students to be college and career ready, students should be exposed to and using the technology available within their schools because technology has become an integral part of our 21st century. King (2012) recommends computer-based assessment because it allows the students to "better demonstrate what they know and are able to do" (p. 5). Additionally, King stated that the advantages of computer-based assessments for teachers and administration would be instantaneous results, which would allow for instructional planning.

New York State is the "pilot" state for implementing computer-based assessments regarding high-stakes testing. King (2012) stated: "The transition to computer-based testing will place New York at the forefront of innovative, 21st century assessment design and delivery" (p. 5). The New York State Education Department plans to implement computer-based assessments within the public schools in April of 2014 (Slentz, 2012).

This assessment change will impact the literacy curriculum and teachers within grades 3-12, who may need to rethink how they deliver assessments within their classrooms. Teachers may consider utilizing computers, and other forms of technology,

as methods of collecting assessment data in addition to the traditional paper-and-pencil assessment method. One point to consider is that teachers will need the technology available within the classroom or school in order to expose the students to testing on computers. Another point to consider is that teachers need adequate training on how to create and administer tests online using computer software, or some other form of online testing.

Considering the literacy curriculum in response to computer-based instruction, classroom teachers may feel the need to modify their instruction. For instance, classroom teachers may incorporate more technology into daily lessons through the use of computers, iPads, or other forms of technology in order to increase students' achievement on computer-based assessments. The switch to computer-based assessments will not only affect teachers, but also the students who will be taking the tests next year. Current students are accustomed to taking New York State Tests within their own classrooms and marking their answers in test booklets using a pencil. Many factors could affect student performance such as the testing environment, familiarity and competence with technology, typing skills, testing accommodations, and cognitive ability.

The purpose of this research was to examine students' performance on computer-based assessments in comparison to paper-based assessments. Additionally, this researcher sought to explore which testing method students preferred. The questions that guided this research were as follows:

❖ How do students perform on computer-based tests versus paper-based tests?

❖ Which testing method do students prefer?

Background

Previous research has examined the use of technology as an instructional tool within a literacy curriculum, but limited studies have looked at using technology as a *primary* vehicle for assessment (Mills & Levido, 2011). The use of paper-based assessment dates back to the 1840s, when Horace Mann utilized written assessments as a form of liability for schools in Massachusetts. During the 1920s, paper-based standardized assessments were widespread and used more frequently as forms of accountability. Schools adopted these assessments in order to receive educational funding. Popularity for standardized, multiple-choice assessments grew in the 1960s where they have been used in school settings since that time (Pearson, Vyas, Sensale, & Kim, 2001). Giving rise to the use of standardized testing from the 1920s and thereafter, the use of computer "hardware" was incorporated in 1955 with the invention of the electronic scanner. The scanner had the ability to correct multiple-choice tests in a timely and cost efficient manner. Testing companies such as National Computer Systems, founded in the 1960s, and Scantron, founded in the 1970s, also utilized the scanner as a way to score large numbers of bubble answer sheets. Starting in the late 1980s, the use of computers changed from scoring standardized tests to administering the tests themselves. For instance, the Graduate Record Examination (GRE) developed a computer-based test in 1992; however, students still had the option to take the paper-based GRE. Today, the GRE is only offered via computer (Clarke, Madaus, Horn, & Ramos, 2000). Currently, there are a limited number of computer-based tests available, and tests that are offered need to be taken on an individual basis. Very few computer-based tests are available to test entire populations of students. However, one such program that does exist is known

as *i-Ready*. Although there are very few states that have implemented computer-based testing for standardized tests, New York State will be one of the first states to implement computerized testing (King, 2012).

In the context of this project, the researcher defines the term *computer-based assessment* to mean that the assessment content is presented digitally using a form of technology such as a computer. The term *paper-based assessment* refers to the assessment whose content is presented in the traditional paper-and-pencil format.

Rationale

The issue of computer-based assessment is significant because implementation of this assessment type is coming very soon. Public school educators in New York State will be administering computer-based assessments for all New York State and Regents exams in approximately a year's time. Furthermore, this issue raises important questions in the field of literacy because it may change the way teachers assess students' knowledge of content within the classroom.

Majority of the studies on computer-based testing versus paper-based testing have used undergraduate students or adolescent students as the participants (e.g., Anakwe, 2008; Bodmann & Robinson, 2004; Clariana & Wallace, 2002; Hardre, Crowson, Xie, & Ly, 2006; Kim & Huynh, 2010; & Li & Pu, 2010). Very little research has been completed in which elementary school students are the participants (e.g., Higgins, Russell, & Hoffman, 2005; Pomplun & Custer, 2005). Therefore, this researcher hopes to fill the gaps in research by assessing elementary school students.

Theoretical Stance

The New Literacies framework (Pacino & Noftle, 2011) correlates with computer-based learning and computer-based assessment. The term "New Literacies" goes beyond the traditional reading of print from a book to involve the reading of digital and multimedia text as well. Students within the current century are experiencing literature through a variety of formats, which include both print and non-print forms. In this digital age, students are being exposed to and growing up with a number of technologies outside of school. These technologies may include, but are certainly not limited to: Xbox Live, PlayStation 3 (PS3), laptops, Global Positioning Systems (GPS), Nooks, Kindles, iPads, iPhones, and smart phones. Although current elementary age students are growing up with technology, there is not enough research that examines their previous experiences with technology and achievement on computer-based tests. Therefore, students who may be "tech-savvy" with their game consoles and kindles may not necessarily perform well on computer-based assessments. According to Pacino and Noftle (2011),

> …Schools, teachers, and parents need to develop a partnership in assisting students to become lifelong learners who engage in critical inquiry as consumers of information/knowledge delivered through many print and non-print formats and who are able to use this information/knowledge to make ethical decisions in a global democracy. (p. 478)

The New Literacies framework correlates with the issue of computer-based testing because New York State is moving toward a digital approach to learning and testing. New Literacies highlight the importance of using a variety of print and non-print

technological literacies to meet the needs of all learners. Teachers who implement

technology into their classrooms will help to prepare students for a world where they can

be college and career ready, as many jobs require using technology in one form or

another. In a 2009 International Reading Association position statement titled, "New

Literacies and 21st Century Technologies," the opening paragraph states:

> To become fully literate in today's world, students must become proficient in the
>
> new literacies of 21st century technologies. As a result, literacy educators have a
>
> responsibility to effectively integrate these new technologies into the curriculum,
>
> preparing students for the literacy future they deserve. (p.1)

This research project aligns with the New Literacies framework because the researcher

examined the use of technology to administer reading comprehension assessments. In

addition, the New Literacies framework includes the use of digital learning and

assessment, and the researcher exposed the students to reading passages via computer.

Lastly, within the New Literacies framework, this researcher explored the opinions of

fourth grade students regarding computer-based testing.

Chapter Two Literature Review

This chapter contains an extensive review of up-to-date empirical research that examines performances and preferences on computer-based tests and paper-based tests. The comparison of scores for computer-based tests and paper-based tests are discussed first, followed by students' preference for each testing format, and then students' cognitive workload in relation to student achievement. Major categories discussed within the literature review include: Similar scores between computer-based tests and paper-based tests, higher results on computer-based tests, higher results on paper-based tests, preference for computer-based tests, and cognitive workload in relation to computer-based testing.

This researcher uses the terms *computer-based test* and *paper-based test* throughout the entirety of this project, however in this chapter, the researcher will keep the terms that were used within the original studies. These terms include: *Paper-and-pencil test, paper-and-pencil assessment, online testing*, *internet test,* and *web-based assessment*.

Comparison of Scores for Computer-Based Tests and Paper-Based Tests

After examining literature related to computer and paper-based assessments, this researcher identified opposing, contradictory findings present within the studies. In several studies, the test scores between computer-based tests and paper-based tests were similar (e.g., Anakwe, 2008; Bodmann & Robinson, 2004; Eid, 2005; Escudier, Newton, Cox, Reynolds, & Odell, 2011; Higgins, Russell, & Hoffmann, 2005; Kim & Huynh, 2010; Li & Pu, 2010). Contradictory findings resulted in students achieving higher scores on computer-based tests (e.g., Clariana & Wallace, 2002; Kapes, Matrinez, & Ip,

1998; Lee, Osborne, & Carpenter, 2010), whereas other studies resulted in students performing better on paper-based tests (e.g., Flowers, Kim, Lewis, & Davis, 2011; Pomplun, Ritchie, & Custer, 2006). Additionally, the participants within each study varied. Some studies (e.g., Anakwe, 2008; Bodmann & Robinson; Clariana & Wallace, 2002; Hadre, Crownson, Xie, & Ly, 2006; Lee, Osborne & Carpenter; 2010) consisted of undergraduate students, while other studies (e.g., Eid, 2005; Flowers, Kim, Lewis, & Davis, 2011; Kapes, Martinez, & Ip, 1998; Kim & Huynh, 2010; Li & Pu, 2010) consisted of high school students as the participants. There were a limited number of studies where elementary-age students were the participants (e.g., Higgins, Russell, & Hoffman, 2005; Pomplun, Ritchie, & Custer, 2006).

Similar scores between computer-based test and paper-based test. Anakwe (2008) compared results of online testing versus paper-based testing. Participants in this study were 75 undergraduate students enrolled in accounting courses. Each student was administered two computer-based tests, and two paper-based tests over the course of one semester. Analysis of the results indicated that there were no significant differences in the students' test scores between the online tests and the paper-and-pencil tests.

Similar to Anakwe (2008), Bodmann and Robinson's (2004) study also used undergraduate students as the participants. Their study included 55 undergraduate students enrolled in an educational psychology course. The students were split into two groups and both groups took a computer-based test and a paper-based test. Parallel to Anakwe's results, the results from this study indicated that there were no score differences between the two administered tests.

In Escudier, Newton, Cox, Reynolds, and Odell's (2011) study, 266 undergraduate dental students were administered online assessments and traditional paper-based assessments.   The content of the exams is unknown, as the researchers simply stated that the students took "related tests of the same knowledge base." The researchers examined the students' performance and perceptions of each testing type. The findings demonstrated that the students' performance yielded similar results for the two types of tests. Additionally, the students perceived both types of test "fair and acceptable."

Other research, which resulted in similar test scores for computer-based tests and paper-based tests, utilized adolescents as the participants. Kim and Huynh (2010) analyzed the ninth grade statewide English test in two testing formats: Paper-based and computer-based. Both tests consisted of 52 multiple-choice questions. The two tests were exactly the same, however the tests were administered to each student in a paper-based format or a computer-based format. Overall, students in the paper-based group and the computer-based group performed at about the same level.

Li and Pu (2010) also investigated the equivalence between computer-based testing and paper-based testing in regards to reading comprehension. Forty-two students, ages 13-16, were split equally into two groups. Each group took a computer-based test and a paper based test. The analysis of the results from the two test types indicated that the scores on both tests were similar.

Eid (2005) administered a computer-based and a paper-based test for mathematics. The tests primarily assessed adolescent students' mathematical problem-

solving skills. The results indicated that the high school students achieved similar scores

on both test types.

A limited number of studies have been conducted in which elementary-aged

students were the participants in relation to computer-based and paper-based testing.

However, Higgins, Russell, and Hoffmann (2005) tested fourth graders' reading

comprehension using three test modes. Students were administered one of the three test

options: A paper-based test, a computer-based test with a scrolling option, or a computer-

based test with a page option. The researchers found no significant score differences.

One possible reason for similar results between computer-based assessments and

paper-based assessments is that most of the participants had previous experiences with

computers. Anawake (2008) mentioned within her study that all of the participants had

taken at least one computer-based test in another undergraduate course. Bodmann and

Robinson's (2004) study described that the participants had previous experiences with

computer-based tests and were familiar with the computer interface that was used within

the study. In Higgins, Russell, and Hoffman's (2005) study, only 9 out of 219 students

stated that they did not have a computer at home (p. 31).

Another possible reason for the similar scores between the two testing types is

because the content of the test was the same. All of the studies mentioned above tested

the students in a particular subject area. If the students knew the content (accounting,

psychology, etc.) then the students should achieve similar test results despite the format

in which the test is presented and administered.

Higher results on computer-based test. Clariana and Wallace (2002) assessed 105 undergraduate students enrolled in a computer fundamentals course. Fifty-four students were given the computer-based test, while the other 51 students were administered the paper-and-pencil test. The results from this study indicated that students performed better on the computer-based assessment than on the paper-and-pencil assessment. Analysis of Clariana and Wallace's study suggests that a possible reason for higher results on the computer-based test is due to the nature of the topics covered in the course. The title of the course was Computer Fundamentals, therefore topics such as computer storage devices, types of printers, resolution of monitors, computer peripheral devices, and types of modems (p. 596) were learned in class. Since the students were using computers as part of the normal curriculum, they were more likely to achieve better results on the computer-based assessment due to their exposure to and familiarity with computers.

Lee, Osborne, and Carpenter's (2010) study examined thirty-one undergraduate students with disabilities as they took part in comparative testing. All undergraduate students who participated were classified with Attention Deficit Hyperactivity Disorder. The students were randomly split into one of four groups: paper-pencil/regular time, paper-pencil/extended time, computerized/regular time, or computerized/extended time. The researchers found that students performed better on the computer-based test as compared to the paper-based versions. They also concluded that extended time versus regular time for the tests did not significantly impact the students' scores. Although, the students who received extended time performed *slightly* better than those with regular time on both testing types.

A study completed by Kapes, Martinez, and Ip (1998) also compared a paper-pencil version of a test to an internet version of a test. The participants included 279 11th and 12th grade students enrolled in Child Care vocational programs and Auto Body vocational programs. Each participate took two tests of the same type. The results showed that higher scores were achieved on the internet test during the second administration. One possibility for higher results on the computer-based tests may be due to a "practice" mode, as the students in each group took the same test twice (Kapes et al., p. 217).

Higher results on paper-based test. Flowers, Kim, Lewis, and Davis (2011) investigated test scores from students with disabilities who had a read aloud accommodation. Scores were analyzed for third through twelfth grade students on reading, math, and science tests. Overall, students performed better on the paper-and-pencil assessment. One possible explanation for higher results on the paper-based test was because students who took the paper-based test were given a read-aloud accommodation, while students who took the computer-based assessment did not receive an accommodation.

In another study conducted by Pomplun, Ritchie, and Custer (2006), test scores were analyzed for students in Kindergarten through third grade to determine if they performed better on a paper-and-pencil test or a computer-based test. More specifically, Pomplun et al. (2006) examined the following domains in relation to literacy: Vocabulary, reading comprehension, and spelling. The results showed that the students earned higher scores on the paper-and-pencil test over the computer-based test. This study also analyzed the factors associated with the results and indicated that the students'

response style as well as the students' socioeconomic status, had an impact upon the results.

Student Preferences on the Two Types of Tests

Some research has examined students' preference regarding computer-based testing and paper-based testing (Escudier, Newton, Cox, Reynolds, & Odell, 2011; Flowers, Kim, Lewis, & Davis, 2011; Hadre, Crowson, Xie, & Ly, 2006; Higgins, Russell, & Hoffmann, 2005; Yurdabakan & Uzunkavak, 2012; Zacharis, 2011). Within these studies, participants were polled through surveys, interviews, or questionnaires and asked which test they preferred to take: Computer-based or paper-based. Within some of the studies, researchers were investigating whether or not the students' test preference would affect their final score.

Preference for computer-based test. In a study previously mentioned, Flowers, Kim, Lewis, and Davis (2011) administered a survey to gain perspective into students' preferences for paper-based testing and computer-based testing. Findings from their survey include a high preference for computer-based testing. However, there was a negative correlation between student preference and test score. The third through twelfth grade students who preferred computer-based testing received lower test scores on that test. Overall, students preferred to take the computer-based test but performed better on the paper-based test.

Higgins, Russell, and Hoffmann (2005) intended to know which type of test fourth grade students preferred taking: Paper-and-pencil, computer test with a scrolling option, or a computer test with a paging option. Eighty-seven percent of 161 fourth graders indicated that they preferred to take the computer-based version of the test rather

than the paper-based version. Lastly, 82% of 135 students stated that it was "easier" to take the test on the computer (p. 32). Although a large percentage of students indicated that they preferred the computer-based test, there were no significant score differences between the three types of tests.

Yurdabakan and Uzunkavak (2012) analyzed third through fifth grade students' attitudes towards computer-based testing. There were 784 students from both private and public schools involved in the research. Findings indicated that students were favorable of computer-based testing. Furthermore, the results implied that students enrolled in public schools favored computer-based testing more so than students enrolled in private schools.

Undergraduate students in Escudier, Newton, Cox, Reynolds, and Odell (2011) deemed online testing as "fair and acceptable" (p. 440) compared with paper-based testing. These students favored the online test compared to the paper-based test. In addition, the results from both types of tests were similar.

Only one study was found in which students favored the paper-based test over the computer-based test. Hardre, Crowson, Xie, and Ly (2006) analyzed differential effects of computer-based, web-based, and paper-based administration of questionnaires. In this study, two groups of undergraduate students were administered three types of questionnaires: Computer-based, web-based, and paper-based. All questionnaires included 16 items. Results suggested that students favored the paper-based questionnaire over the computer-based and web-based questionnaires.

Based upon a review of the literature, the overall theme was that students prefered computer-based testing over paper-based testing (e.g., Escudier, Newton, Cox, Reynolds,

& Odell, 2011; Flowers, Kim, Lewis, & Davis, 2011; Higgins, Russell, & Hoffmann, 2005; Yurdabakan & Uzunkavak, 2012). Very few studies demonstrated that students prefer to take a paper-based test over a computer-based test (e.g., Hardre, Crowson, Xie, and Ly, 2006).

Cognitive Workload in Relation to Student Achievement

In addition to comparing students' scores on computer-based tests versus paper-based tests, researchers have also examined students' cognitive workload in regards to their scores (e.g., Liu & Su, 2011; Noyes, Garland, & Robbins, 2004; Stevenson, Touw, & Resing, 2011). The analyses from these three studies demonstrate that there is an increased amount of effort on the part of the learner when faced with computer-based learning and assessments. However, studies reached opposing conclusions in relation to assessment results.

Li and Su (2011) suggested that students' cognitive workload was higher in a computer-based learning environment, and that the students performed *better* within the computer-based environment. Liu and Su (2011) conducted a study to compare high school students' cognitive workload in a multimedia learning environment with computer simulations to a traditional learning environment with teacher lectures and demonstrations. Although the computer-based learning demanded more on the cognitive load of the learner, the students within the computer simulation group outperformed the traditional group. One possible explanation for students performing better within the computer-based learning environment was because the students were actively involved in the learning process and were required to create meaning from the computer simulations (p. 606).

According to Noyes, Garland, and Robbins (2004), the "task demands" of the computer-based assessment may be too difficult for younger students to manage (p. 111). Younger students, kindergarten through third grade, may experience cognitive overload when testing on the computer. The researchers stated that students' cognitive workload was high when exposed to a computer-based assessment, which caused the students to perform better on the paper-based test. Overall, the researchers concluded that "more effort appears to be needed to complete a computer-based test" (p. 112).

Stevenson, Touw, and Resing (2011) also concluded that the cognitive workload demands of students increase when testing on the computer. However, they found similar results between the two testing types. The researchers found that computer-based testing was a suitable assessment method for five-year olds, as the results demonstrated similar scores between the computer-based test and the paper-based test. The researchers suggested that the computer-based test should "mimic" the paper-based test as much as possible to ensure accuracy of scores in comparison to the paper-based test.

Although the studies reviewed all concluded that there are high cognitive demands placed upon students when faced with a computer-based test, each study demonstrated different results from computer-based tests versus paper-based tests. One study concluded that students performed better on computer-based tests (Liu & Su, 2011), another study concluded that students performed better on the paper-based tests (Noyes, Garland, & Robbins, 2004), and yet another study concluded that students performed similarly on the computer-based test versus the paper-based test (Stevenson, Touw, & Resing, 2011).

Coping with the Assessment Change

Since the switch from paper-based assessments to computer-based assessments is mandated for New York State public schools, teachers will have to cope with this change. Teachers can cope with this change in three ways: By understanding that elementary students need to become proficient in computer skills such as keyboarding, using a mouse, and navigating a webpage; by incorporating technology into lesson daily plans; and by using technology to deliver assessments.

There is a limited body of research relating to elementary school students and their computer proficiency. Very few studies exist in which researchers have examined the *effects* of using technology in correlation with elementary students' learning. Furthermore, there are a limited number of studies that discuss the implementation of technology instruction *and* student learning outcomes within the elementary grades.

Horne, Ferrier, Singleton, and Read (2011) examined 952 students, from ages 11-17, and their handwriting skills versus their typing skills. During the study, the students handwrote a passage that was dictated from a computer, and typed another passage that was also dictated from a computer. The passages were about 200 words in length. The results demonstrated that the older students (age 14-17) outperformed the younger students (age 11) with typing skills. The 11-year old students typed an average of 14 words per minute and could handwrite faster than they could type. The older students, ages 14-17, typed an average of 27 words per minute and could type faster than they could handwrite. One possible explanation for the results is that the older students had more experiences with typing, which is why they were able to type more words per minute than the younger students. It was observed by the researchers that as the students

increased in age, their typing skills became more proficient and the students were able to type faster than they could handwrite.

Another study completed by Kemker, Barron, and Harmes (2007) examined the use of laptops as a form of authentic learning in a low-income fifth grade classroom. The researchers observed twenty-six students and their teacher during a two-year period. The students had little or no access to technology prior to the start of the project. Each student was given a laptop computer to utilize throughout the school day. The students participated in a number of activities and projects in which the students were held accountable to complete, with modeling and guidance from the teacher. Programs such as word processors, spreadsheets, graphic organizers, video editing software, digital cameras, and scanners were used on a daily basis. The researchers found that the students were highly motivated to complete lessons that required the use of technology, as they set their own learning goals. The researchers stated: "Students learned to use the computer as tools to learning, much as they would a pencil or paper" (p. 312). To sum up, the laptops were a "tool" used by the students to complete authentic lessons which reflected real world applications.

The current body of research related to computer-based assessments suggests that there is a high demand placed upon the cognitive workload of students when faced with a computer-based test. Yet, the research demonstrates differing results in regards to the two testing types. Computer-based assessments have been used with varying degrees of success, with older participants. Additionally, the current research suggests that older students prefer computer-based assessments over paper-based assessments. In order to understand the effects of using computer-based assessments with the elementary

population, further research should be conducted on the cognitive demands, score

differences, and preferences in relation to the two testing types within elementary

classrooms.

Chapter Three Methods

The purpose of conducting this research was to explore how fourth grade students with and without disabilities performed on computer-based tests versus paper-based tests. This study also intended to investigate students' preference for each testing method. Data included: Results from computer-based and paper-based assessments, student surveys, observational notes, and interviews with students.

Participants

Within the researcher's fourth grade classroom, seven out of eleven students participated in the study. The fourth grade students attended a small, rural Western New York school which had a total of 362 students enrolled in pre-kindergarten through sixth grade. A large majority of the students who attended the school were considered to have low socioeconomic status. Furthermore, the school district was considered a low-income school, as many of the students qualified for free or reduced lunches. The participants in the study included one boy and six girls. The pseudonym for the male student was Ethan, who was ten years old. The pseudonyms for girls who participated in the study were: Kate, Carol, Michelle, Jill, and Brook, all of whom were nine years old except Ashley, who was ten years old. All of the seven students were Caucasian, and the primary language in which they communicated was English.

Ethan, Brook, and Ashley had an Individualized Education Plan (IEP) and were classified with a Speech/Language Disorder or Learning Disability pertaining to reading and/or mathematics. Traditional testing accommodations for these students during a paper-based test included: Extended testing time, separate location, directions to be read aloud, and students to restate directions to ensure understanding. All of the participants

had previous experience with weekly computer lab lessons, as well as monthly iPad lessons within the classroom. Additionally, the students had taken a limited number of computer-based assessments. These assessments were completed using the i-Ready software program, which tested the students' knowledge of reading comprehension, vocabulary, phonemic awareness, and phonics, as well as mathematical concepts and skills.

Within the classroom setting Jill, Kate, and Carol were all considered proficient readers. In general, each of the three girls earned high scores on the weekly reading tests. In comparison, Ethan and Brook were less proficient readers, and were considered struggling readers. They consistently earned low test scores on the weekly reading tests. Ashley and Michelle were considered average readers and often earned average scores on the weekly reading tests.

Data Collection

Data were collected on Tuesday afternoons from March 12, 2013 – March 19, 2013 at the research lab during the regularly scheduled computer lab time. Data sources included: Results from two computer-based comprehension tests and two paper-based comprehension tests, two student surveys which included five questions regarding the testing experience and opinions of the assessments, open-ended interviews with students in small groups, and the researcher's observational notes. The seven fourth grade students were randomly split into two groups: A and B. Group A consisted of four students, while group B had three.

Test format. The content of the reading comprehension assessment, whether computer-based or paper-based, was exactly the same. Both tests contained a reading

passage with an average of 494 words, five multiple-choice questions with four choices, and two short response questions. The multiple-choice questions were worth a total of two points each and there was one correct answer of four choices. The short response questions were worth three points each. The participants earned one point for restating the question, one point for answering the question correctly, and one point for correct grammar and punctuation. Altogether the tests were worth a total of 16 points.

The computer-based assessment was created by the researcher and was administered to the students through the *Toolbox Pro* software. On the computer-based assessment, the students read the passage on the first screen. After reading the passage, the students clicked an arrow, which took them to a new screen, and the passage was no longer available. On the next screen, the students read a multiple-choice question and selected an answer by clicking on a circle, which represented their choice: A, B, C or D. The students saw only one multiple-choice question at a time. They had to select an answer before moving onto the next question and click an arrow to direct them to the next screen. Once they selected an answer, they could not go back to the previous page to change an answer. Lastly, the students typed answers for two short response questions using the keyboard. One short response question appeared at a time. The students clicked the *submit* button once they completed the computer-based assessment. Students were not able to make any changes to the computer-based test. Furthermore, students did not have the ability to modify the screen, go back to the passage while answering the test questions, or change an answer once a selection was submitted.

On the paper-based assessment, the students read the passage on page one and answered five multiple-choice questions by circling letter A, B, C, or D, on the page.

The students handwrote answers for two, short response questions on the paper-based assessment. Additionally, the students had the opportunity to go back and change their answers, if they so desired.

Testing accommodations for students with IEPs included: Extended testing time, directions read aloud, and students to restate directions to ensure understanding. Students that required testing accommodations received them for the computer-based test as well as the paper-based test.

The researcher took observational notes while the students completed the assessments. The researcher's observational notes included the following items: Students' behavior during the assessment, students' comments during the assessment, and the length it took each student to complete the assessment.

In addition to the researcher's observational notes and the student surveys, the students participated in small group interviews with the researcher about the assessment that was previously administered. The focus of the interview was on the students' experience with each testing method, their preference or non-preference for the testing method, and their opinions of the assessment. The interviews were recorded using a digital voice recorder. The time frame for the assessment, surveys, and interviews was a total of 45 minutes per day.

Test administration. On March 12, 2013, all students took the first reading assessment in the research lab. Group A took the first reading comprehension assessment on the laptop computers within the research lab, while Group B took the same reading comprehension assessment using the paper-and-pencil method. Each student took the paper-based survey as he/she completed the assessment. When all of the students in

Group A were finished with their computer-based assessment and their survey, the researcher took the students to an empty classroom to conduct a brief, audio-recorded interview. The students in Group B stayed in the research lab with the computer lab teacher and had a quiet reflection time. At the conclusion of the interview with Group A, the students returned to the research lab and stayed with the computer lab teacher, while the researcher took Group B to the empty classroom to conduct a brief, audio-recorded interview. Upon completion of the interview with Group B, the researcher returned to the research lab and gathered all students to resume the day's lessons and activities.

On March 19, 2013 both groups took the second reading assessment at the research lab. Group B took the second reading comprehension assessment on the laptop computers within the computer lab; Group A took the same reading comprehension assessment using the paper-and-pencil method. Again, each student took the paper-based survey as he/she completed the assessment. When all of the students in Group B were finished with their computer-based assessment and their survey, the researcher took the students to an empty classroom to conduct a brief, audio-recorded interview. The students in Group A stayed in the research lab with the computer lab teacher and had a quiet reflection time. At the conclusion of the interview with Group B, the students returned to the research lab and stayed with the computer lab teacher, while the researcher took Group A to the empty classroom to conduct a brief, audio-recorded interview. Upon completion of the interview with Group A, the researcher returned to the research lab and gathered all of the students to resume the day's lessons and activities.

Data Analysis

       After the data were collected, the researcher analyzed: the students' scores on each of the four tests, responses from the student surveys, the transcribed interviews with the students, and the observational notes.

Computer-based tests and paper-based tests. After each student completed a computer-based test and a paper-based test, the researcher analyzed the results from the two tests. The goal of this comparison was to discover a trend of students' performance comparing computer-based tests and paper-based tests.

Student surveys. The researcher examined the surveys and coded them according to perceived level of difficulty, whether or not the students went back to change their answers on the paper-based tests, and if the students liked the testing methods. The purpose was to identify the pattern for student preference.

Audio-taped interviews. The audio-taped interviews were transcribed by the researcher, typed into word documents, and then coded according to commonalities regarding test preference and accommodations. Again, the codings included the following: Perceived level of difficulty, perceived length of the assessment, whether or not the students went back to change their answers on the paper-based tests, if the students liked the testing methods, and changes that the students would make to the assessments. The surveys and interviews allowed the researcher to identify the pattern of students' preference on testing format.

Observational notes. The researcher's observational notes were also coded according to key categories that emerged from the data. When the researcher analyzed all of the data collectively, the researcher intended to find out how fourth grade students, with and

without disabilities, performed on computer-based tests and paper-based tests. The

researcher gained insight into the fourth graders' opinions and preferences regarding the

two testing methods. Lastly, the researcher examined the correlation between the test

scores and the students' preferences.

Role of the researcher/classroom teacher. This researcher explored whether the students

in her fourth grade classroom performed better on one testing method over another:

Computer-based versus paper-based. Additionally, the researcher was interested to find

out which testing method the students preferred. She did not allow her role as a

classroom teacher to interfere with the judgments she made as a researcher.

In order to diminish coercion, this researcher took necessary steps to ensure that

results of the research were valid and reliable. The students were not given any sort of

reward for participating in the study. The results of the tests were used for research

purposes only and were not included in the students' report card grades. Additionally,

when the fourth graders granted consent to participate in the study, the researcher was not

present in the classroom. She left the classroom and another fourth grade teacher read the

consent form to the students. After the fourth grade teacher read the consent form aloud,

the students either granted consent to participate in the study or denied consent to

participate. The teacher collected the forms from the students and then the researcher

came back into the classroom.

The researcher sent home parent/guardian consent forms which described the purpose of

the study, explained that the results of the study would not correlate with their report card

grades, ensured that all students' names would be changed to protect the real identity of

the students, explained that the students could withdraw from the study at any time, and

ensured that the students' privacy would be accounted for at all times. Lastly, when the

assessments were being delivered, the researcher maintained her role as a classroom

teacher by providing accommodations to those students who had them indicated on their

IEPs. Other than that, the researcher did not probe the students, nor offered them any

kind of assistance throughout the entirety of the assessment, either computer-based or

paper-based.

Timeline

The table below represents the dates and events for the research project.

Tuesday, March 12, 2013	❖ Group A took the computer-based test ❖ Group B took the paper-based test ❖ All students completed a survey after the test. ❖ The researcher conducted an interview with Group A and B after completion of the surveys.
Tuesday, March 19, 2013	❖ Group A took the paper-based test ❖ Group B took the computer-based test ❖ All students completed a survey after the test. ❖ The researcher conducted an interview with Group A and B after completion of the surveys.

IRA Matrix

The following table displays how this research project met the International

Reading Association's 2010 Standards for Reading Professionals. Through this action

research project, the Reading Specialist/Literacy Coach Candidate met the standards in

the following ways:

Table 1
IRA Matrix

IRA Standard	How the Thesis Met the Standard
Element 1.1 – Foundational Knowledge *Candidates understand major theories and empirical research that describe the cognitive, linguistic, motivational, and sociocultural foundations of reading and writing development, processes, and components, including word recognition, language comprehension, strategic knowledge, and reading–writing connections.*	Within Chapter One, specifically the theoretical framework section,' the Literacy Coach Candidate discusses how this research project correlates to the New Literacies framework.
Element 2.1 – Curriculum and Instruction *Candidates use foundational knowledge to design or implement an integrated, comprehensive, and balanced curriculum.*	The Literacy Coach Candidate develops and implements paper-based and computer-based assessments, which correlate to the literacy curriculum. This research project will shed light on the needs of all students through the examination of their testing accommodations.
Element 2.3 – Curriculum and Instruction *Candidates use a wide range of texts (e.g., narrative, expository, and poetry) from traditional print, digital, and online resources.*	By utilizing both print and online reading passages for testing, the Literacy Coach Candidate demonstrates her understanding that a variety of print, digital, and online formats should be utilized within the curriculum and instruction.
Element 3.1 – Assessment and Evaluation *Candidates understand types of assessments and their purposes, strengths, and limitations.*	Within chapter two, the Literacy Coach Candidate thoroughly demonstrates an understanding of the literature and research related to assessments and their uses and misuses.
Element 3.2 - Assessment and Evaluation *Candidates select, develop, administer, and interpret assessments, both traditional print and electronic, for specific purposes.*	This entire research project correlates with this standard. The Literacy Coach Candidate develops and administers paper-based and computer-based assessments to fourth grade students. Thereafter, the Literacy Coach Candidate interprets the assessments for the purpose of comparing results between the two testing methods.

IRA Standard	How the Thesis Met the Standard
Element 3.4 – Assessment and Evaluation *Candidates communicate assessment results and implications to a variety of audiences.*	During the professional development workshop, the Literacy Candidate will provide support for teachers in the analysis of data, using the assessment results of all students. Additionally, the Literacy Candidate will lead teachers in analyzing and using classroom, individual, grade-level, and schoolwide assessment data to make instructional decisions.
Element 4.2 – Diversity *Candidates use a literacy curriculum and engage in instructional practices that positively impact students' knowledge, beliefs, and engagement with the features of diversity.*	During the professional development workshop, the Literacy Coach Candidate will support classroom teachers to provide differentiated instruction and instructional materials; including traditional print, digital, and online resources, all of which capitalize on diversity.
Element 5.3 – Literate Environment *Candidates use routines to support reading and writing instruction (e.g., time allocation, transitions from one activity to another, discussions, and peer feedback).*	The Literacy Coach Candidate will administer the assessments during the regularly scheduled computer lab time, so as not to interfere with the normal classroom routines.
Element 6.3 – Professional Learning and Leadership *Candidates participate in, design, facilitate, lead, and evaluate effective and differentiated professional development programs.*	Through the professional development workshop, the Literacy Coach Candidate is collaborating in planning, leading, and evaluating professional development activities for individuals and groups of teachers. Support will be given to teachers to boost their efforts to use technology in literacy assessment and instruction.

Chapter Four Results and Interpretation

Results from computer-based tests and paper-based tests, observational notes, student surveys, and interviews were analyzed in order to identify general patterns and themes related to computer-based and paper-based testing as well as students' preference for each testing method. First, each type of data was analyzed and interpreted individually and categories emerged from each data source. Then, all four data types were interpreted collectively. General patterns are discussed within this chapter.

Test Results

Paper-based tests. Carol, Brook, and Jill took the first paper-based test on March 12, 2013. The average score for this set of tests was a 77%. The highest score was a 94% and the lowest score was a 56%.   Michelle, Ethan, Kate, and Ashley took the second paper-based test on March 19, 2013. The average for this set of tests was a 68%. The highest score was a 100% and the lowest score was a 38%.

Computer-based tests.   Michelle, Ethan, Kate, and Ashley took the first computer-based test on March 12, 2013. The average score for these tests was a 49%. The highest score was a 75% while the lowest score was a 25%. Carol, Brook, and Jill took the second computer-based test on March 19, 2013. The average score for this set of tests was a 32%. The highest test score was a 44% and the lowest test score was a 13%.

Student Surveys

The results from the surveys indicated that the majority of students liked the paper-based test more so than the computer-based test.   All of the students, except for Ethan, indicated their preference for the paper-based tests. When asked if the students *liked* the computer-based test, three students said "yes;" while four students said "no."

Interviews

All students were interviewed after their tests; therefore each student was interviewed twice. Similar to the survey results, the majority of students preferred the paper-based test to the computer-based test. During the first and second round of tests all of the students, except for Ethan, stated that they preferred the paper-based test.

Observational Notes

Four patterns emerged from the observational notes. Most of the students looked back within the passage on the paper-based test. This indicated that students employed looking back as a way to help them answer comprehension questions. Further, students utilized tools and strategies while they completed the paper-based test. Students used their pencils to trace, underline, cross out, and circle important words within the passage or test questions. In addition, students demonstrated a lack of computer skills; specifically with keyboarding. Several of the students asked questions that specifically related to the computer-based test.

Interpretation of Data

Two major themes emerged across the various data sources. The majority of students performed better on the paper-based tests, except for Michelle (see Figure 1). Additionally, most of the students preferred the paper-based test. Therefore, their preference positively correlated with their overall performance. Two possible explanations for higher results on the paper-based tests were: Students used test-taking strategies on the paper-based test and the students were less experienced with testing on the computer.
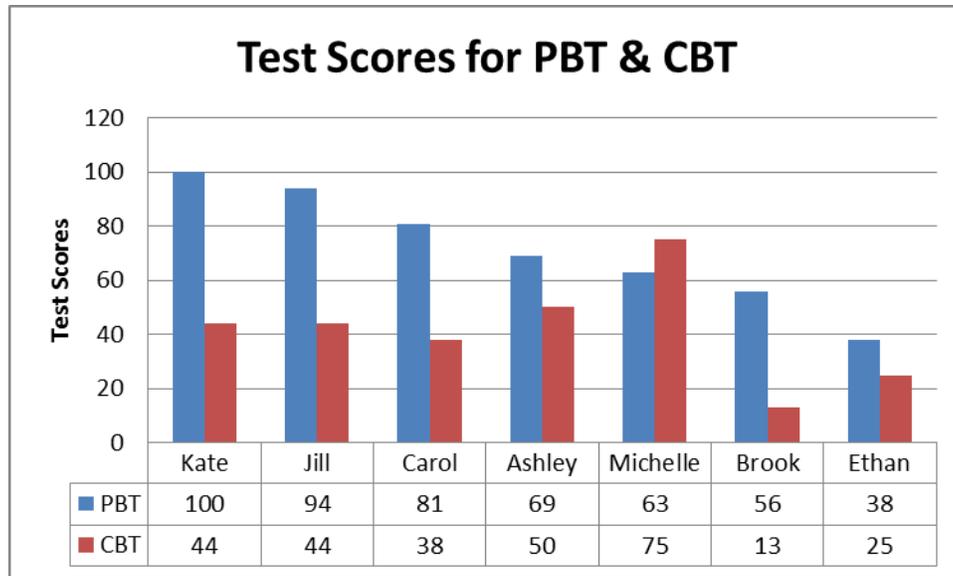
Test Scores for PBT & CBT

	Kate	Jill	Carol	Ashley	Michelle	Brook	Ethan
PBT	100	94	81	69	63	56	38
CBT	44	44	38	50	75	13	25

Figure 1

Test-taking strategies. The first possible reason for higher scores on the paper-based test was because the students employed strategies as they completed the test. Several of the data sources demonstrated that the students looked back at the passage, and that students used tools. It was indicated from the observational notes that five of the seven students looked back in the passage as they answered the test questions on the paper-based test. In addition, on the survey, five of the seven students stated that they changed one or more answers before they turned in their paper-based test.

Another pattern from the observational notes was the students' use of tools such as erasers, pencils, and highlighters to help them on the test. Students used their pencils to trace, underline, cross out and circle important words within the passage or the test questions. The observational notes identified four of the seven students who used their pencils to mark up their paper-based test.

Furthermore, students who scored significantly higher on the paper-based test employed more strategies as they took the paper-based test. It can be seen in Figure 2

that Kate, Jill, Carol, and Brook performed considerably higher on the paper-based test,

and the tests, observational notes, student surveys, and interviews indicated that these

students utilized more test-taking strategies than the other students.

Test Scores for PBT & CBT

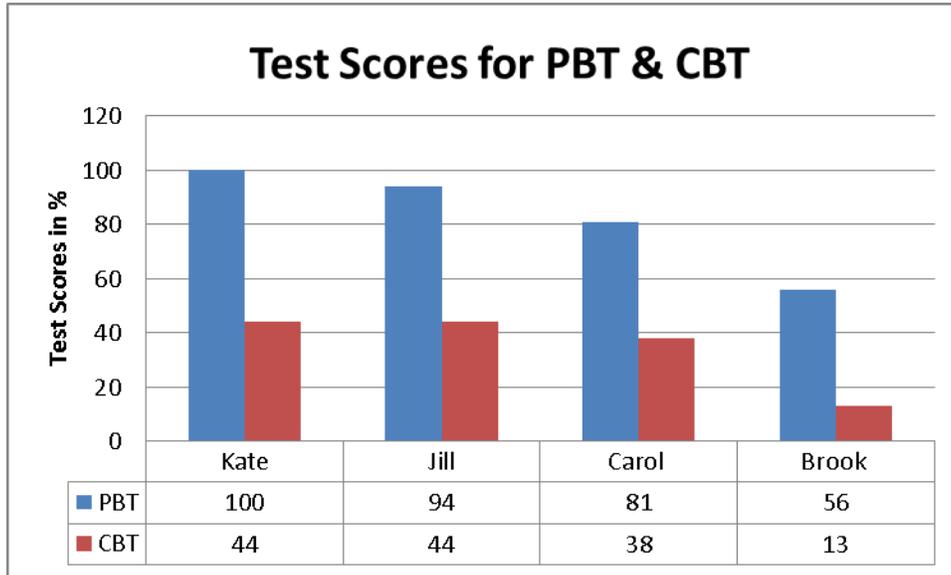|	Kate	Jill	Carol	Brook
PBT	100	94	81	56
CBT	44	44	38	13

Figure 2

As indicated in the observational notes, all four of the students used the passage to

help them answer the test questions. It was noted that Kate used her pencil to trace the

words as she read the passage. She also circled vocabulary words within the passage and

circled important words and phrases within the test questions. Additionally, Kate and Jill

crossed out letters on the multiple-choice section to eliminate answers. On the short

response questions, Brook underlined text which she used to restate the question within

her answers. She received full credit on the paper-based test for the short response

questions. However, on the computer-based test, she did not receive any credit for the

short response questions. This proves that Brook's use of tools helped her to be more

successful on the paper-based test.

On the student surveys, both Carol and Kate stated: "I like this testing method (paper-based) because you can go back and change it, on the computer you cannot." On her computer-based survey, Kate stated: "I do not like this testing method because you cannot go back to the passage." Lastly, Carol, Kate, and Brook indicated on the student survey that they went back and changed one or more answers before they turned in their test.

Ashley, Michelle, and Ethan performed about the same on both of the tests, as there was not a significant score difference from their paper-based test to their computer-based test. This can be explained by the students' lack of testing strategies on the paper-based test. Unlike the four students who earned considerably higher scores on the paper-based test, Ashley and Michelle did not use any tools to help them on the test. They did not underline, circle, or trace any words within the passage or test questions. On her survey, Ashley stated that she did not change any of her answers before she turned in her paper-based test. Ethan employed one strategy on the paper-based test. He crossed out multiple-choice answers to eliminate choices, but he performed about the same on both tests.

Less experience with computer-based testing. Students' low scores on the computer-based test can be explained due to their lack of experience with testing on the computer. Two patterns that emerged from the observational notes support that the students lacked experience with computer-based testing. Students did not demonstrate proper keyboarding and several students asked questions as they completed the test. It was evident that five of the seven students lacked computer proficient skills, specifically

with typing. These students typed using their index fingers only. Carol was seen typing

with her hand in the formation of an arch, as if she were holding a baseball.

It was also noted that students asked questions as they completed the computer-

based test. Michelle and Jill, on two different testing days asked: "Can I make this

bigger?" Their question specifically referred to the size of the text on the screen. Kate

came across a problem as she completed her computer-based test. At one point, her

entire computer screen went red. She was unsure of what to do and asked the researcher

for help. The observational notes also indicated that Brook did not know what to do after

she typed her short response answer. She sat there for a while and the researcher asked

her: "Are you done with that one? Then click the green arrow to move onto the next

question."

Preference

The overall consensus was that the majority of students preferred the paper-based

test to the computer-based test. Ethan was the only student who indicated a preference

for the computer-based test. During the interview, Ethan stated that he preferred the

computer-based test because he could type.

One explanation for the students' preference for the paper-based test can be

related to the students' lack of experience with testing on computers. As previously

mentioned, it was evident on the paper-based test that the strategies the students learned

in the classroom were utilized when they took the test. The students knew how to use

their tools (e.g., pencil, eraser, and highlighter) to help them on the paper-based test. On

the computer-based test, the students did not have enough experience with tools (e.g.,

mouse, highlight function, and underline function) and therefore did not know how to use

tools during the computer-based test. Jill made a comment during an interview, which conveyed that she preferred the paper-based test due to her use of physical tools. She stated: "I like the paper-based test because you can highlight stuff and you can see what stuff is important." During the same interview, Carol stated: "I like the paper-based test because you can go back and change the answers." The researcher asked her: "What tool can you use to do that?" Carol answered: "Your eraser." Brook also commented on the use of tools. She stated: "I prefer the paper because you can read it and underline words."

The majority of students who preferred the paper-based test more so than the computer-based test earned a higher score on the paper-based test. Their preference for the paper-based test positively correlated with their actual performance. One outlier was in Ethan's case. Several data sources demonstrated that Ethan preferred the computer-based test. On the survey, Ethan indicated that he did not like the paper-based test because "I would rather take it on the computer." On another survey he wrote: "I like the computer-based test because it had typing involved." During the interview, he stated: "I like computer because you can type and the reason I don't like paper is because you have to write." Although Ethan preferred the computer-based test to the paper-based test, he performed slightly lower on the computer-based test.   He earned a 38% on the paper-based test and a 25% on the computer-based test.

Reliability of Data

Data sources within this study were reliable for a number of reasons. First of all, the tests themselves measured the students' comprehension of a text. The test did not measure anything else but comprehension. Reliability was assured, as the researcher did

not simply gather data from a single source. There were multiple sources of data within

this study such as test scores, surveys, interviews, and observational notes, to triangulate

the findings. Lastly, the researcher was assured that the students gave their actual

opinions within the surveys and interviews because the researcher did not offer them any

reward or consequence for their opinion, and they had nothing to lose or gain. Due to the

fact that nothing was offered to them (positively or negatively), the researcher believed

the students were honest.

Chapter Five Discussion and Conclusion

Overview of Study and Findings

This study was an action research project that centered on students' performance on different formats of reading assessments. Seven fourth grade students took a paper-based test and a computer-based test, completed two surveys, and participated in two interviews. Factors were identified that impacted fourth graders' performance on paper-based tests and computer-based reading comprehension tests. Two major themes emerged from the data: 1. Students performed higher on the paper-based test; and 2. Students preferred the paper-based test.

Significance of the Findings

This study added to the current research that focused on computer-based assessments. The results from this study were important because New York State will implement computer-based testing by April of 2014. Findings from this study showed that students, specifically fourth graders, were not yet ready for computer-based assessments. It was evident that the students from this study lacked experience with testing on the computer as well as computer proficient skills, specifically keyboarding. Additionally, testing strategies that were learned in the classroom were not applicable for the computer-based test. Based upon the results from this study, the Commissioner of Education should reconsider implementation of computer-based testing for elementary-age students.

Limitations of the Study

While this study contributes significantly to the current field of literacy research on assessment, there were various limitations. One limitation was the sample size. This

research was conducted in one school within one classroom that had seven out of 11 students who participated in the study. This study also examined one age group. This study did not analyze in great depth the students' socioeconomic status nor gender. Those factors may have also contributed to the difference in test scores. Due to the small sample size, the results from this study should not be generalized.

Another limitation was the number of tests administered and the subject in which the test was administered. Each student took one paper-based test and one computer-based test. Future studies could include administering more than one test of each type to see if there is a general pattern in the test scores. Additionally, studies in other subject areas should be completed to see if the results are consistent with this study.

It was not clear whether the format or content of the tests affected the students' performance. Within this study, the format was the same, however the researcher did not specifically examine the format. Moreover, the content of the tests was not explored in relation to performance. Students' background knowledge of the content was not taken into consideration when test scores were analyzed.

Finally, cognitive workload was not measured within this study, but was examined in the literature review section of this project. Future studies on elementary-aged students' cognitive workload in relation to computer-based testing should be completed.

Conclusion

Results from this study demonstrated that the majority of participating fourth grade students earned higher scores on the paper-based reading comprehension test. Overall, the students preferred the paper-based test to the computer-based test. Therefore

their preference for the paper-based test impacted their overall performance on the tests. Since computer-based testing is proposed to take effect in April of 2014, classroom teachers should expose students to computer-based testing. Based upon the findings from this study, it is assumed that more experience with computer-based testing will result in test results that appropriately reflect students' actual proficiency in the subject area. A final implication for teachers is that students need to be taught test-taking strategies for computer-based testing. This study demonstrated that students performed significantly higher on the paper-based test because they utilized test-taking strategies. Therefore, students should learn testing strategies in order to perform adequately on computer-based tests.

Recommendations for Future Research

While this study adds to the current research on assessment, further research should be conducted on this topic, as there are not many studies that focus on elementary-age students as the participants. Additional studies should be done on paper-based testing versus computer-based testing, with elementary-age participants, to see if there is a consistent trend with the findings of this study.

Another recommendation for future studies is to examine how student preference impacts performance on computer-based and paper-based testing. There are a limited number of studies that have examined this topic.

Future research should also examine other subject areas in relation to paper-based testing versus computer-based testing. Tests in other subject areas should be included to determine if the content or format impacts the students' overall score. This study did not solely focus on content or format, but considered them as the tests were designed for the

study. Therefore, a final topic that could be explored is the content versus the format of

the tests.

References

Anakwe, B. (2008). Comparison of student performance in paper-based versus

Computer-based testing. *Journal of Education for Business, September/October,*

13-17.

Bodmann, S. M., & Robinson, D. H. (2004). Speed and performance differences among

Computer-based and paper-pencil tests. *Journal of Educational Computing*

Research, 31(1), 51-60.

Bridgemann, B., Lennon, M. L., & Jackenthal, A. (2003). Effects of screen size, screen

resolution and display rate on computer-based test performance. *Applied*

Measurements in Education, 16, 191-205.

Bromley, K. (2010). Picture a world without pens, pencils, and paper: the unanticipated

future of reading and writing. *Journal of College Reading and Learning, 41*(1),

97-108.

Clariana, R., & Wallace, P. (2002). Paper-based versus computer-based assessment: key

factors associated with the test mode effect. *British Journal of Educational*

Technology, 30(5), 593-602.

Clarke, M. M., Madaus, G. F., Horn, C. L., & Ramos, M. A. (2000). Retrospective on

educational testing and assessment in the 20th century. *Journal of Curriculum*

Studies, 32(2), 159-181.

Escudier, M. P., Newton, T. J., Cox, M. J., Reynolds, P. A., & Odell, E. W. (2011).

University students' attainment and perceptions of computer delivered

assessment; a comparison between computer-based and traditional tests in a

'high-stakes' examination. *Journal of Computer Assisted Learning, 27*, 440-447.

Eid, G.K. (2005). An investigation into the effects and factors influencing computer-
    based online math problem-solving in primary schools. *Journal of Educational
    Technology Systems, 33*(3), 223-240.

Flowers, C., Do-Hong, K., Lewis, P., & Davis, V. C. (2011). A comparison of
    computer-based testing and pencil-and-paper testing for students with a read-
    aloud accommodation. *Journal of Special Education Technology, 26*(1), 1-12.

Harde, P. L., Crowson, H. M., Xie, K., & Ly, C. (2007). Testing differential effects of
    computer-based, web-based, and paper-based administration of questionnaire
    research instruments. *British Journal of Educational Technology, 30*(1), 5-22.
    Doi: 10.1111/j.1467-8535.2006.00591.x

Higgins, J., Russell, M., & Hoffmann, T. (2005). Examining the effect of computer-based
    Passage presentation on reading test performance. *The Journal of Technology,
    Learning, and Assessment, 3*(4), 1-36.

Horne, J., Ferrier, J., Singleton, C., & Read, C. (2011). Computerised assessment of
    handwriting and typing speed. *Educational & Child Psychology, 28*(2), 52-66.

Kapes, J.T., Matinez, L., & Ip, C.F. (1998). Internet-based vs. paper-pencil occupational
    competency test administration: an equivalency study. *Journal of Vocational
    Education Research, 23*(3), 201-219.

Kemker, K., Barron, A., & Harmes, J.C. (2007). Laptop computers in the elementary
    classroom: authentic instruction with at-risk students. *Educational Media
    International, 44*(4), 305-321.

Kim, D. H., & Huynh, H. (2010). Equivalence of paper-and pencil and online

administration modes of the statewide English test for students with

and without disabilities. *Educational Assessment, 15*, 107-121.

Doi: 10.11080/10627197.2010.491066

King, J. B. (2012). New York state's transition to computer-based testing. *The State*

Education Department Memo to Superintendents, Principles, Professional

Organizations. Retrieved from www.p.12.nysed.gov/transtion_strategy_ to

computer_based_testing.

Lee, K.S., Osborne, R.E., & Carpenter, D.N. (2006). Testing accommodations for

university students with AD/HD: computerized vs. paper-pencil/regular vs.

extended time. *Journal of Education Computing Research, 42*(4), 443-458.

Li, M., & Pu, H. (2010). Comparison between CBT and PBT: assessment of gap-filling

and multiple-choice cloze in reading comprehension. *Journal of Language*

Teaching and Research, 1(6), 935-941. Doi: 10.4304/jltr.1.6.935-941

Mills, K. A. & Levido, A. (2011). IPed: pedagogy for digital text production. *The*

Reading Teacher, 65(1), 80-91.

Noyles, J., Garland, K., & Robbins, L. (2004). Paper-based versus computer-based

assessment: is workload another test mode affect? *British Journal of Educational*

Technology, 35(1), 111-113.

Pearson, D. P., Vyas, S., Sensale, L. M., & Kim, Y. (2001). Making our way through the

Assessment and accountability maze: Where do we go now? *The Clearing House,*

74(4), 175-182.

Pacino, M. A. & Noftle, J. T. (2011). New literacies for global, digital learners. *The International Journal of Learning, 18*(1), 477-485.

Pomplun, M., Ritchie, T., & Custer, M. (2006). Factors in paper-and-pencil and computer reading score differences at the primary grades. *Educational Assessment, 11*(2), 127-143.

Russell, M. (1999). Testing on computers: A follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives, 7*, 1-20.

Slentz, K. (2012). Common core sample questions from new york and parcc. *The State Education Department Memo to Superintendents, Principles, Professional Organizations.* Received August 17, 2012 via e-mail.

Stevenson, C. E., Touw, K. W. J., & Resing, W. C. M. (2011). Computer or paper analogy puzzles: Does assessment mode influence young children's strategy progression? *Educational & Child Psychology, 28*(2), 67-84. Yurdabakan, I. & Uzunkavak, C. (2012). Primary school students' attitudes towards computer based testing and assessment in turkey. *Turkish Online Journal of Distance Education, 13*(3), 177-188.

Appendix

Appendix A

Survey and Interview Questions

Paper-Based Survey Questions:

1. I like this testing method   yes / no
   a. Explain why or why not.
2. Paper-test only: Did you go back and change any of your answers? yes / no
3. Computer-test only: Would you like to go back and change any of your answers? yes/no
4. If I could change one thing about this test it would be: _____
5. Please rank the following things to help you with taking the test:
   I'd like more time
   I'd like someone to read the questions aloud to me
   I'd like the passage to be shorter in length
   I'd like to answer fewer questions
   I'd like to take the test in the classroom

Interview Questions:

1. What do you think about the taking the test on computer/paper?
2. How do you think you did on this test?
3. Which testing method do you prefer? Explain why.