Assessing online assessments: A comparison study of math assessment tools for third-grade students

Tina Kim Chan

Follow this and additional works at: https://scholarworks.lib.csusb.edu/etd-project

Part of the Science and Mathematics Education Commons

Recommended Citation
https://scholarworks.lib.csusb.edu/etd-project/3074
ASSESSING ONLINE ASSESSMENTS: A COMPARISON STUDY OF
MATH ASSESSMENT TOOLS FOR THIRD-GRADE STUDENTS

A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts
in
Education:
Instructional Technology

by
Tina Kim Chan
March 2007
ASSESSING ONLINE ASSESSMENTS: A COMPARISON STUDY OF
MATH ASSESSMENT TOOLS FOR THIRD-GRADE STUDENTS

A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

by
Tina Kim Chan
March 2007

Approved by:

Dr. Eun-Ok Baek, First Reader

Dr. Brian Newberry, Second Reader
ABSTRACT

The study reported here examined the move towards online assessments and addressed the question of whether or not different assessment tools affect student scores and student learning.

The research activities covered a three-week period, from June 5, 2006 to June 23, 2006. During this time, seventeen third grade students served as their own control group by taking several math tests online and several math tests with paper and pencil. Results were compared to see if performance on computer-based tests would be more successful than pencil-and-paper tests.

Findings revealed that 59% of the students did better on the paper tests and 41% of the students did better on the computer tests. Further analysis revealed that the overall average of the computer tests was 71.9%, while paper tests revealed an average of 69.9%. In summary, these findings report that there is no significant difference in scores when taking a test on the computer or a test on paper.
ACKNOWLEDGMENTS

I would like to thank my first reader, Dr. Eun-Ok Baek for taking the time to guide me through this research study. Your passion for Instructional Technology and academics has highly motivated me in completing my thesis. I’d also like to thank my second reader, Dr. Brian Newberry. Your help and suggestions have been priceless.
DEDICATION

This book is dedicated to the four most influential people in my life. To my Heavenly Father, You are my strength when I am weak. Your unconditional love and care have been with me all my life. To my parents, you’ve taught me to pursue my dreams despite the hardships. Thank you for bringing me to America and giving me opportunities to further my academics. Finally, to my loving husband, Matthew. By believing in me even before I believed in myself, you’ve given me the courage to accomplish my goals and to pursue my dreams. Without your wonderful sense of humor I would not have gotten through these last few torturous years of graduate school. I love you all.
# TABLE OF CONTENTS

ABSTRACT ............................................ iii
ACKNOWLEDGMENTS .................................... iv
LIST OF TABLES ...................................... vii
LIST OF FIGURES ..................................... viii

CHAPTER ONE: BACKGROUND

Introduction ....................................... 1
Statement of the Problem ........................... 3
Purpose of the Study ............................... 4
Research Questions and Research Hypotheses ... 5
Significance of the Research ....................... 6
Limitations of Study Design and Procedures ...... 7
Definition of Terms ............................... 8

CHAPTER TWO: REVIEW OF THE LITERATURE

Introduction ....................................... 10
Current Trends in Computer-based Testing ....... 12
Analysis of Computer-based Testing in the Market ......................................... 15
Advantages and Challenges of Computer-based Testing ..................................... 17
Comparison Between Computer-based Testing and Paper-based Testing ............... 19
Implementation of Computer-based Tests ........ 25
Summary ........................................ 26

CHAPTER THREE: METHODOLOGY

Introduction ....................................... 28
## Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Sections</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td></td>
<td>29</td>
</tr>
<tr>
<td>Research Design and Instrument Development</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>Reliability and Validity</td>
<td></td>
<td>32</td>
</tr>
<tr>
<td>Data Analysis</td>
<td></td>
<td>39</td>
</tr>
<tr>
<td>Summary</td>
<td></td>
<td>40</td>
</tr>
<tr>
<td><strong>CHAPTER FOUR: RESULTS AND DISCUSSION</strong></td>
<td></td>
<td>41</td>
</tr>
<tr>
<td>Introduction</td>
<td></td>
<td>41</td>
</tr>
<tr>
<td>Presentation of the Findings</td>
<td></td>
<td>41</td>
</tr>
<tr>
<td>Student Math Test Results</td>
<td></td>
<td>41</td>
</tr>
<tr>
<td>Student Survey Results</td>
<td></td>
<td>44</td>
</tr>
<tr>
<td>Discussion of the Findings</td>
<td></td>
<td>50</td>
</tr>
<tr>
<td>Summary</td>
<td></td>
<td>55</td>
</tr>
<tr>
<td><strong>CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS</strong></td>
<td></td>
<td>56</td>
</tr>
<tr>
<td>Introduction</td>
<td></td>
<td>56</td>
</tr>
<tr>
<td>Conclusions</td>
<td></td>
<td>56</td>
</tr>
<tr>
<td>Research Questions and Hypothesis Conclusions</td>
<td></td>
<td>56</td>
</tr>
<tr>
<td>Meaning of Study</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>Future Research and Recommendations</td>
<td></td>
<td>61</td>
</tr>
<tr>
<td>Summary</td>
<td></td>
<td>63</td>
</tr>
<tr>
<td><strong>APPENDIX A: PARENT INFORMED CONSENT</strong></td>
<td></td>
<td>64</td>
</tr>
<tr>
<td><strong>APPENDIX B: STUDENT ASSENT FORM</strong></td>
<td></td>
<td>66</td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
<td>68</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1. Key D'Arcy Factors Compared to County and State ...................................... 29

Table 2. Number of Students Scoring in the Four State-generated Categories ............. 30

Table 3. Student Questionnaire Form ................................................................. 34

Table 4. Overall Mean Scores for Computer-based Test versus Paper-and-pencil Test ........... 42

Table 5. Results of Question One ..................... 45

Table 6. Results of Question Two ..................... 46

Table 7. Results of Question Three ..................... 47

Table 8. Results of Question Four ..................... 48

Table 9. Results of Question Five ..................... 49
LIST OF FIGURES

Figure 1. Kids' Place Houghton Mifflin Mathematics Website ................................................. 36
Figure 2. Paper-and-pencil Version Assessment ........ 37
Figure 3. Online Website ........................................ 38
Figure 4. Sample Online Test Question ................. 39
Figure 5. Comparison of Means for Each Test Taken .... 43
Figure 6. Comparison of Boys and Girls for Paper and Computer Test .............................. 44
CHAPTER ONE
BACKGROUND

Introduction

With the overwhelming amount of paperwork that teachers come into contact with in one day, most teachers would agree that there needs to be a more efficient way of grading papers. The more traditional and long-established way of grading papers usually consisted of a red pen, a basic calculator, and a grade book that had numerous lines constituting rows and columns for student names, assignment headings, and averages. Teachers would spend countless hours hand-grading assignments and tests, writing down scores in tiny little boxes, and eventually, at the end of each grading period, meticulously and tediously calculate averages for grades.

The moment I began my teaching career, a part of me rebelled against this time-honored tradition of paper-and-pencil grading. I found it not only time consuming, but also a hazard environmentally; the amount of paper that is consumed daily by the average student is alarming. As a result of my rebellion, I became interested in finding a more efficient, effective and valuable way of grading papers. Online assessments and online grading
seemed to be the perfect match to my dilemma. As time went on, it seemed only natural that my passions for efficient grading lead me to research this topic.

The focus on assessment has increased due to the No Child Left Behind Act (Neugent, 2004). The No Child Left Behind Act requires states to test students more often to prove they are meeting national achievement standards. As a result, online assessment is one of the fastest growing segments in education publishing (Billings, 2004). Educators, administrators and districts are now turning to computer-based tests to get a more accurate view of a student’s knowledge and abilities.

Online assessments are computer-based tests that automatically score student tests. Most include extensive reporting features and suggestions for remediation (Billings, 2004). These reporting features can include student’s academic weak areas and student’s academic strengths. It may also give teachers key areas to focus on with instructional materials on missed exam questions, resource materials, and additional standards-aligned questions.

Online assessments are also known as computer-based testing, computerized testing, electronic testing and computer adaptive testing. Traditional paper-and-pencil
tests do not allow schools and educators to get immediate feedback. According to Alexander Russo, computer-based testing can provide flexibility, instant feedback, individualized assessment and eventually lower costs than traditional paper examinations (2002, pp. 6-12).

Statement of the Problem

Over the years educators have found many ways to assess their students’ learning. Traditionally, a student’s accountability involved teachers using a variety of assessment tools, such as homework, quizzes, oral presentations, and tests. Over the past few years, there has been more of an emphasis on computer-based testing as a means to measure student learning. As such, the question of equivalence has been frequently asked among those using online assessments.

The move towards online assessments has raised many questions to whether different assessment tools affect student scores and student learning. More specifically, educators are asking whether online testing scores are equivalent to the traditional paper-and-pencil testing scores. In other words, if a student took similar tests on the computer and on paper, would scores be equivalent?
Purpose of the Study

Currently, the education community is not fully clear to whether different assessment tools affect student’s test performance. A previous study was done where freshman business undergraduates were randomly assigned either to a computer-based or identical paper-based test. Data from this study showed that the computer-based test group outperformed the paper-based test group (Clariana, 2002). On the other hand, Olsen, Maynes, Slawson, and Ho’s study showed that no significant difference existed between paper-administered tests and computer-administered tests (1986).

The purpose of this study is to see if third-grade students who take the math tests online will score higher than if they took the math tests with the traditional medium. Another purpose is to understand students’ thoughts regarding online assessments and if given the choice would they prefer to take computer tests rather than paper-and-pencil tests. Furthermore, the question of gender is an important factor to high test scores. This research study will also look at the role gender plays in computer-based testing.

Besides being a relatively new field, school districts have had positive experiences to online testing.
School districts are finding online testing to be easy to use and valuable to students and teachers alike. Proponents predict that computerized testing will be widespread within just a few years (Russo, 2002).

Research Questions and Research Hypotheses

The study began with three specific research questions:

- Will students score higher on the computer test than the paper-and-pencil version?
- If students were given the choice, would they prefer to take tests on the computer rather than the more traditional medium?
- Does gender play a role in how well students succeed on the computer tests?

The study permitted examination of the following hypotheses:

- There will not be a difference in scores between students who take the math tests on the computer and those who take it on paper.
- If given the choice, students would not prefer to take tests on the computer rather than the more traditional medium.
• The difference in gender does not play a role in how well students succeed on computerized tests.

Significance of the Research

The increased testing requirements for the No Child Left Behind Act has forced educators to find an alternative way to evaluate students. Computerized tests can improve accuracy in both the taking and scoring of tests. Computerized testing allows students to take exams in which only one question is asked at a time. This lessens the possibility of filling out an answer sheet incorrectly. Furthermore, computerized tests theoretically yield 100% accuracy of results. With computerized tests meeting today's tech-savvy generation of kids, ever-faster, more accurate and clearer ways of testing will continue to develop in the future. The timeliness and relevance of this research can not be overlooked. More and more districts are using computer-based testing to fulfill their requirements with the state. This research will give more knowledge into whether computerized tests are equivalent with paper-and-pencil tests. As more research is conducted on the subject matter, the use of computerized testing will be more spread in areas of education.
Limitations of Study Design and Procedures

Several limitations of this study must be noted. First, this study included a small sample size of third-grade students. This might have limited its generalizability. Colorado State University defines generalizability in the following way (2006): "In many ways, generalizability amounts to nothing more than making predictions based on a recurring experience. If something occurs frequently, we expect that it will continue to do so in the future." In my study only seventeen students participated in the study. A wider sampling of students in different classrooms, schools, or districts may have provided a clearer representation of computer scores and paper scores. Secondly, by increasing the number of tests taken and math problems answered by students, results might have been different. Because the study was conducted over a short period of time, students were only given the chance to take fifteen online tests and fifteen paper-and-pencil tests. Each of these tests had only eight questions, thus giving a total of 120 math problems for computer and 120 math problems for paper-and-pencil. Finally, the math questions that were given to the students were related to third-grade math standards. Students had just been taught that school year the
third-grade math standards. Students would have been more comfortable answering math review questions from the second grade.

Definition of Terms

The following terms are defined as they apply to the project.

**Computer-based testing:** Also known as online testing, electronic testing, computer adaptive testing, and web-based testing. Testing is programmed and then administered to students on computer. Question formats are frequently objective, with discrete-point items and these tests are subsequently scored electronically.

**Paper-and-pencil testing:** Also known as traditional testing. Testing is administered to students on paper and answered with a pencil or pen. These tests are subsequently scored by teacher.

**Equivalence:** Equivalence exists when paper-and-pencil test can be translated into a computer-based assessment form and scores from the two tests are interchangeable.

**Low-to-no stakes assessment:** These assessments do not have high stakes like the end-of-the-year state tests.
These short and frequent tests pinpoint student’s academic weaknesses and allow teachers to teach accordingly to that student’s needs.

**Computer anxiety**: Computer anxiety refers to the fear experienced when interacting with a computer or anticipating an interaction (McDonald, 2002).

**Computer attitudes**: Computer attitudes overlap with computer familiarity and computer anxiety. In other words, computer attitudes are typically shaped by computer experience and computer anxiety or confidence. The importance of positive attitudes towards computers has been considered a prerequisite for developing computer-skills. It is important to recognize that individual differences have the potential to significantly affect the equivalence of computer and paper-based assessments (McDonald, 2002).
CHAPTER TWO

REVIEW OF THE LITERATURE

Introduction

Computer-based assessment programs have been made increasingly available. Many instructors are beginning to translate traditional paper-and-pencil tests to electronic formats using these very easy-to-use programs. These assessment programs allow educators to create and customize educational software online that is built around their own course materials. The online tests are then made available to students over the Web. Quia Web (located at http://www.quia.com/web) is one of the world’s most popular educational technology Web sites. Quia Web is designed to make an educator’s job easier, save teachers more time, help instructors motivate and engage students, and help students learn more and learn faster.

The switch from paper-and-pencil tests to computer-based tests can also be seen in a variety of exams, including state drivers’ licensure, military training, job applications in the private sector, postsecondary education, and professional certification (Russo, 2002; Trotter, 2001). The move towards online assessments has also occurred in the primary grades.
The No Child Left Behind Act of 2001, commonly known as NCLB, is a United States federal law that allows for a number of federal programs to improve the performance of United States' primary and secondary schools. The law increases the accountability for states, school districts and schools by setting high expectations, goals, and standards. The No Child Left Behind Act of 2001, championed by the Bush administration, calls for all public-school students to be proficient in reading and math by 2014. Schools must make steady progress towards these goals. They face penalties if they don't continually raise their proportion of proficient students, both overall and within various racial and other categories. Schools that miss milestones can be required to pay for outside tutors and let parents transfer children elsewhere (Golden, 2003). As a result of the No Child Left Behind legislation, there has been an increased interest in computer-based assessments (Trautman, 2004). A school needs to perform and make "Adequate Yearly Progress" in order to receive proper funding and support. Administrators and educators are forced to find an alternative way to quickly and successfully identify student's weak areas and accordingly provide remediation for those weak areas. Computer-based assessments seem to
provide schools with the right tools to facilitate student learning.

This chapter will examine the recent development of computer-based assessments and the phenomenon and trend of online assessments. It will also review current research related to the comparison of computer-based assessments to paper-and-pencil assessments.

Current Trends in Computer-based Testing

Computer-based testing is also known as online assessment, computerized testing, electronic testing and computer adaptive testing (Russo, 2002). All these computer-based programs have one thing in common: they enable educators and students to get immediate feedback. Some programs also allow for detailed reports of student’s scores, which in turn allow for teachers to adjust instruction to the student’s performance level.

Most include extensive reporting features and suggestions for remediation (Billings, 2004). Traditional paper and pencil tests do not allow schools and educators to get immediate feedback. According to Alexander Russo, "Computer-based testing can provide flexibility, instant feedback, individualized assessment and eventually lower costs than traditional paper examinations (2002,"
pp. 6-12).” He and Tymms also state, “Computer-based testing and computer-assisted assessment have many advantages over traditional paper-and-pencil testing and assessment. These include immediate, unbiased and accurate scoring and feedback, increased efficiency, convenient individualized administration and improved test security (2005, pp. 419-429).”

Maria Northcote (2002) writes that online assessments in the 21st century have come to mean many things to a range of people in various educational contexts. For some educators, online assessments are looked upon as a friend. For others, it is perceived as a foe. Online assessments can provide an avenue for designing authentic, relevant tasks in order to assess student learning outcomes or online assessments is a mechanism with the potential to unnecessarily reduce the quality of learning experienced by online learners from deep to surface levels. Other educators think of it as a fix. Online assessments have the potential to reduce the time and cost commitments associated with administering, distributing and marking high numbers of student assignments and examinations.

Students also have many different views of online assessments. They may perceive it valuable because of immediate feedback from teachers or find it frustrating
because of technological breakdowns and withdrawal
symptoms from the lack of face-to-face contact with
teachers.

Administrators also look upon online assessments in
various ways. Online assessments can provide an efficient
means by which to collect grades or be associated with
plagiarism, equity and the cost of specific software
licenses. Maria Northcote concludes that one thing is for
certain—the increasing use of online assessments (2002).

Karen Billings writes that assessment is one of the
fastest growing segments in education publishing (2004).
She defines online assessments as computer-based tests
that automatically score student tests. Most programs
include extensive reporting features and suggestions for
remediation. Already in the works are computer programs
that can electronically score essays and short-answer
responses (Russo, 2002).

The Software and Information Industry Association
(SIIA) surveyed top computer-based assessment vendors to
learn how they have changed to meet the growing demands
for online assessments. Their responses were basically the
same (2004):

a. Online assessment results need to improve
   instruction.
b. Assessments need to be linked to state standards.

c. Online assessments need to reassess student progress.

These responses imply that the future of online assessments will be used to improve instruction, remediation for students at-risk, and will be lined to state standards.

Analysis of Computer-based Testing in the Market

So far, school districts have had positive experiences to online testing. Proponents predict that computerized testing will be widespread within just a few years (Russo, 2002). A variety of computer software systems have been developed for use in computer-based testing and analysis. These include systems delivered over the Web (e.g. CASTLE, Hot Potatoes, WebMCQ, WebTest, TAL, TOAL, Questionmark™ Perception™, i-assess) for wide access, systems delivered over an organization's local area network (LAN) (e.g. Questionmark™ Perception™, i-assess) and systems installed on individual computers (e.g. FastTEST, MicroCAT, Questionmark™ Perception™) (He, 2005).
There seems to be a consensus regarding the use of online assessments:

- "The move toward online testing enables everyone to get immediate feedback" (Trautman, 2004, p. 8).
- "Online tests are a more accurate view of a student's knowledge and abilities" (Trautman, 2004, p. 8).
- "...increased student motivation" (Russo, 2002, p. 6)
- "...improves instruction" (Billings, 2004, p. 26)
- "Computer-based testing can provide flexibility, instant feedback, individualized assessment and eventually lower costs than traditional paper examinations" (Russo, 2002).
- "Since the assessment data is submitted electronically, it is generally easy for the instructor to collect the data to a spreadsheet
for record keeping or for further analysis" (Sanchis, 2001).

- "One of the most valuable benefits of using computer-assisted assessment in education is that it can provide timely and specific information on the performance of each student which can be used for diagnosing areas where students have individual difficulties" (He, 2005).

Advantages and Challenges of Computer-based Testing

The future seems endless as computer-based assessments begin to evolve into a mandatory testing requirement for all school districts. Catherine McCaslin, the evaluation and assessment specialist for the Beaufort County, S.C., states, "Why not make testing fun? Whoever said that a test had to be boring and quiet and black and white to be a valid classroom assessment tool" (2002, pp. 6)?

Online assessment tools are now being used to improve teaching. Known as low-to-no-stakes assessment, this kind of assessment has no consequence to either the school or the student. The sole purpose of low-to-no stakes assessment is to inform teaching (Trautman, 2004). These
short and frequent tests pinpoint student’s academic weaknesses and allow teachers to teach accordingly to that student’s needs.

It must be noted that continuous assessment of student progress must be provided (Billings, 2004). Computer-based assessments will be most effective when students are allowed the opportunity to be assessed, receive accurate academic instruction by teacher and then further reassessed.

Advocates argue that computerized tests are more appropriate for today’s technology-savvy youngsters (Bushweller, 2000). For the second year, third-graders in Anchorage, Alaska are taking an online reading exam. The two-part test evaluates students’ abilities to read independently. On the multiple-choice part of the exam, questions appear on the screen, students click on what they think is the correct answer, and the computer instantly gives them feedback on their performance. The second part of the tests consists of short essay questions. These questions are also online, but the students give their responses on paper, and the essays are graded by teachers (Bushweller, 2000). Alison Haigler, a third-grade teacher in Anchorage says, "Taking tests is not a fun thing, but for the kids, taking [tests] by
computer is more exciting than getting out a test booklet. They all enjoy working on computers—I think that gets them a little more motivated. If I had had the option to take tests by computer, I would have done it” (2000, p. 3).

The advantages to the test taker are not as clear. Researchers have found that test takers’ attitudes range from enthusiasm at being able to schedule exams when they want and to get results instantly to active, long-term dislike of computers (Bugbee, 1996). Nonetheless, it seems from test takers’ acceptance of computerized testing that the perceived benefits of computerized tests outweigh the perceived benefits of paper-and-pencil tests. However, it is also necessary to examine the effects of online testing on students’ academic achievements and motivational aspects.

Comparison Between Computer-based Testing and Paper-based Testing

Jane Healy (1999), an educational psychologist and author of *Failure to Connect: How Computers Affect Our Children’s Minds and What We Can Do about It*, states that schools should be aware that computerized testing can raise equity issues. She also states that parents and teachers should be very concerned that the conditions under which their children are tested are fair. Using
computers may put some children at a disadvantage and others at an advantage.

Research is not clear whether different assessment tools affect student's test performance. A previous study was done where freshman business undergraduates were randomly assigned either to a computer-based or identical paper-based test. Data from this study showed that the computer-based test group outperformed the paper-based test group (Clariana, 2002). Another study from Sinclair, Renshaw and Taylor suggest that students who took the paper-based test scored higher overall (2004).

On the other hand, Olsen, Maynes, Slawson, and Ho's study showed that no significant difference existed between paper-administered tests and computer-administered tests (1986). Kapes, Martinez, Ip, Slivinski, and Hardwicke's study also showed the same results. In their study, they compared traditional paper-pencil administration of occupational competency tests with the same tests administered over the internet. Their study found no statistically significant differences between the two versions of tests (1998). Another study included twenty-seven introductory psychology students. These students completed ten identical unit tests via paper or computer. The results indicated that, when students were
motivated and testing conditions were equivalent, there were no differences between the scores obtained via computer-based or paper-pencil tests. It is evident that there still is considerable debate concerning the equivalence of scores obtained from different testing tools.

There is a need to establish equivalence between paper-and-pencil test and computer-based tests. If a paper-and-pencil test has been translated into a computer-based assessment form, it is necessary to establish score and construct equivalence if scores from the two tests are to be interchangeable. If school districts decide to go paperless, the knowledge of equivalence will allow for a seamless transition to computer-based testing.

Angus McDonald writes that most studies on score equivalence have largely ignored individual differences such as computer experience, computer anxiety and computer attitudes (2002). The extent to which test takers have experience of using computers has been argued to influence their performance on computerized tests. Taylor, Kirsch, Eignor, and Jamieson (1999) found computer experience to encompass "experience, frequency of use, type of use, number of courses involving computers, owning a computer,
access to computers, attitudes towards computers, and related technologies" (pp. 219-274).

Computer anxiety refers to the fear experienced when interacting with a computer or anticipating an interaction (McDonald, 2002). It has been assumed that computer anxiety results from a lack of familiarity with computers, and so given adequate exposure anxiety should be reduced.

Computer attitudes overlap with computer familiarity and computer anxiety. In other words, computer attitudes are typically shaped by computer experience and computer anxiety or confidence. The importance of positive attitudes towards computers has been considered a prerequisite for developing computer-skills. It is important to recognize that individual differences have the potential to significantly affect the equivalence of computer and paper-based assessments (McDonald, 2002).

An informal survey was given to ask Oregon students what they thought about the state’s computerized assessments. The survey showed students finding computer-based testing faster and more enjoyable than the traditional paper and pencil tests. The students also reported feeling more confident on their performance on computerized assessments than on traditional tests (Park, 2003).
During the pilot phase of the online testing program, Oregon state education officials surveyed 740 third graders and 730 high school students about their experiences taking the state assessment online compared to the hard-copy version. Third graders were especially positive about Web-based testing. Seventy-nine percent of the third graders surveyed believed they had done their best work on the computer-based reading test, compared with a little more than five percent of third graders who felt they had done better on paper or were frustrated using the computerized test. Sixty-two percent reported that the online reading test was easier to use and more enjoyable than the paper test, and fifty-eight percent said the same about the Web-based test in mathematics (Park, 2003).

High school students were slightly less positive about their performance on the Internet-based tests. Thirty-seven percent of the high school students rated the computerized version of the reading test easier to use and more enjoyable to take than the paper version. Thirty-eight percent found the Web-based math test easier to use and more enjoyable to take (Park, 2003).

A majority of the high school students reported that the Internet version had taken less time or the same
amount of time as the paper test. But a considerable number of students disagreed. Thirty-one percent of the high school students tested said the Web version of the reading test had taken more time than the paper test. The same percentage reported that the computerized math test had taken longer to complete than the paper test (Park, 2003).

Another study compared paper-and-pencil and a web-base version of the Ansell-Casey Life Skills Assessment (ACLSA) for mode equivalence (Bressani, 2002). Ninety-seven young people completed an age-appropriate ACLSA version in both formats. No significant main effects between modes of administration emerged. However, higher scores were achieved on the second administration regardless of format. These findings demonstrate the equivalence of Internet and paper-based formats for all three ACLSA levels and a predictable learning effect across the first and second testings (Bressani, 2002).

Author Tom Trautman writes about schools moving towards online assessments (2004). He addresses the issue that the results received from traditional paper-and-pencil tests does not allow schools and educators to react quickly enough to the changes that need
to be made. On the other hand, online testing enables everyone to get immediate feedback.

Implementation of Computer-based Tests

Tom Trautman (2004) concluded with three components to look for with online assessment programs:

a. Look for evaluation tools that can create assessments to address the learning standards for your state.

b. Look for assessment tools that can generate short and frequent evaluations to inform instruction.

c. Look for assessment tools that provide feedback for individual students on individual standards.

Author Lan Neugent addresses three areas of focus when educators get ready for online testing: establishing a project management team, having technological goals, and overcoming challenges (2004). Being aware of these issues will allow administrator, teacher and student become successful with online testing.

For example, the Virginia Department of Education first established a project management team to develop and guide what they called the Web-based Standards of Learning Technology Initiative. Second, their technological goals
were to establish student access to computers at a ratio of one computer for every five students, create Internet-ready LAN capability in every school, and assure adequate high-speed, high-band-width capability. Lastly, the Virginia Department of Education communicated with the schools to address issues promptly so not to lead to local testing disasters.

Lan Neugent concluded by saying, "When you decide to take your tests online, remember to plan well, involve everyone, get the correct technology, and then address the assessment issues." Being aware of details will make the transition from paper-and-pencil testing to computer-based testing all the more easier.

Summary

The increased testing requirements for the No Child Left Behind Act will completely revolutionize online assessments in the 21st century. The current trends in instructional technology suggest that educators are ready and willing to try a new type of assessment tool that can alleviate the burden of high-stakes testing requirements. Educators are excited by the ease and practical nature of these online testing tools. While there are advantages to computer-based testing over traditional testing, there is
still considerable debate concerning the equivalence of scores obtained from these testing modalities (Mason, 2001). This research endeavors to answer the many questions regarding the equivalence of computer-based assessments versus paper-and-pencil testing.
CHAPTER THREE

METHODOLOGY

Introduction

As stated in Chapter 1, the study reported here examined the move towards online assessments and addressed the question of whether or not different assessment tools affect student scores and student learning. Additionally, the study also examined gender and computer experience as variables to high test scores on computer-based assessments.

This chapter is organized in terms of three specific research questions posed in Chapter 1:

- Will students score higher on the computer test than the paper-and-pencil version?
- If students were given the choice, would they prefer to take tests on the computer rather than the more traditional medium?
- Does gender play a role in how well students succeed on the computer tests?

This chapter will endeavor to outline the research design behind online assessments and paper-and-pencil assessments equivalency.
Participants

Using convenient sampling, one third-grade classroom consisting of seventeen students from Michael D'Arcy Elementary in Fontana, California was assessed. D'Arcy is one of eighteen elementary schools in the Colton Joint Unified School District. During the 2003-2004 school year, around 721 students were enrolled at D'Arcy. Of those students, 8.7% were still learning English and 46.5% were considered low-income students. Table 1 shows key factors of D'Arcy compared to county and state. These key factors are shown because they may affect how well students perform on future assessments.

Table 1. Key D'Arcy Factors Compared to County and State

<table>
<thead>
<tr>
<th></th>
<th>Students</th>
<th>ELL</th>
<th>Low Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>D'Arcy</td>
<td>721</td>
<td>63(8.7%)</td>
<td>335(46.5%)</td>
</tr>
<tr>
<td>County</td>
<td>419,084</td>
<td>79,646(19%)</td>
<td>213,733(51%)</td>
</tr>
<tr>
<td>State</td>
<td>6.3 million</td>
<td>1,575,000(25%)</td>
<td>3,087,000(49%)</td>
</tr>
</tbody>
</table>

The seventeen third grade students chosen for this study represented a wide range of academic abilities in mathematics. Some were high math achievers (advanced), others average achievers (proficient) and others low achievers (basic, below basic, or far below basic.) Table
Table 2 shows data from last year's 2004-2005 California Standards Tests in Mathematics.

Table 2. Number of Students Scoring in the Four State-generated Categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Advanced</th>
<th>Proficient</th>
<th>Basic</th>
<th>Below Basic</th>
<th>Far Below Basic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>7 (44%)</td>
<td>4 (25%)</td>
<td>5 (31%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

*One student's scores were not available. Table 2 is based on only 16 students.

In regards to gender, of the seventeen participants, 11 (65%) were females and 6 (35%) were males. These seventeen students represent a range of variables. All students have access to their school computer lab once a week with their classmates. Some have access to computer technology at home, while others do not. For those students who have computer access at home, the time spent and the purpose of computer experiences differ.

The research activities covered a three-week period, from June 5, 2006 to June 23, 2006. During this time, these third grade students served as their own control group by taking several math tests online and several math tests with paper and pencil. Results were compared to see
if performance on computer-based tests would be more successful than pencil-and-paper tests.

Research Design and Instrument Development

This research began with a quantitative approach as the primary method, and then used a follow-up survey to evaluate and interpret the quantitative results. In this quasi-experimental research, students’ online scores were compared to students’ own traditional paper-and-pencil scores. This time-series design was based on the collection of quantitative observations at regular intervals through repeated assessments. In other words, students took an 8 question online test, and then alternated with a different 8 question traditional test. These math test items were related to the third-grade state standards the students were currently learning in their classroom. Most test questions were related to mathematical reasoning, number sense, measurement, geometry, statistics and data analysis. Ultimately, each student completed a total of 240 questions. Therefore fifteen tests were taken online and fifteen tests were taken with the traditional paper-and-pencil method. In this manner, students served as their own control group and thus had scores that would be comparable.
Reliability and Validity

Reliability is the extent to which an experiment, test, or any measuring procedure yields the same result on repeated trials (Colostate, 2005). To ensure reliability in this research, the instruments for testing need to be accurate. Additionally, third-grade review items were used for testing. The test items were not over new material thus eliminating the variable that some of the test questions were too hard or never taught. Finally, to ensure that the two testing instruments were equivalent the exact paper test was transcribed word-for-word over to the computer test.

Validity refers to the degree to which a study accurately reflects or assesses the specific concept that the researcher is attempting to measure (Colostate, 2005). In reference to the two main instruments used, the test items and the student survey question, both instruments showed accurate soundness. The test questions were taken from a very reliable website (http://www.eduplace.com). Houghton Mifflin was chosen because of the superior quality of educational materials that are published year after year. For more than 150 years, Houghton Mifflin has been one of the country’s premier educational publishers. Houghton Mifflin provides educational materials, trade and
reference books, and assessments. The math questions that were used in this study proved equal to the superior quality of all educational materials that Houghton Mifflin produces each year. Each math question that was used in the study accurately assessed the math knowledge third-graders are required to know in third grade. Furthermore, in order to assure validity, the student’s survey items were reviewed by two experts to measure students’ preference correctly.

The second part of the research included a survey. This survey was given to the third-graders after they had completed the math tests. The survey included five questions directly related to their past computer experience and also their attitudes towards taking the math assessments online and on paper and pencil. Table 3 shows the survey given to students after the testing sessions.
Table 3. Student Questionnaire Form

1. How often do you use computers?
   a) I use a computer often, both at home and at school.
   b) I use a computer sometimes, mostly at home.
   c) I use a computer sometimes, mostly at school.
   d) I never or almost never use a computer.

2. How do you feel about using computers?
   a) I feel I do my best work on a computer.
   b) I am comfortable using a computer.
   c) Using computers does not appeal to me.
   d) I am frequently frustrated when I use a computer.

3. How would you feel about taking a state test on a computer?
   a) It would be easier and more enjoyable than a paper and pencil test.
   b) It would be about the same as taking a paper and pencil test.
   c) It would be harder than taking a paper and pencil test.
   d) It would be more difficult at first, but become easier as I got used to it.

4. On the computer part of the test
   a) I did my best work.
   b) I didn’t do as well as I did on the paper and pencil part of the test.
   c) I did equally well on the computer and paper and pencil parts of the test.
   d) I was so frustrated using the computer; I was unable to do my best work.

5. The time needed to take the computer part of the test was...
   a) less than a paper and pencil test would have taken.
   b) more than a paper and pencil test.
   c) about the same as a paper and pencil test.
   d) I didn’t notice how long the computer test was taking.

Before any research began, an application was submitted to the Institutional Review Board Committee at
California State University, San Bernardino. Parents and students understood the research goals, the risks, if any, and benefits (see Appendix A and Appendix B).

Both versions of the math tests had questions taken from the Kids’ Place Houghton Mifflin Mathematics’ website (http://www.eduplace.com/kids/mhm/testprep/gr3/index.html). Houghton Mifflin was chosen because of the superior quality of educational materials that are published year after year. For more than 150 years, Houghton Mifflin has been one of the country’s premier educational publishers. Houghton Mifflin provides educational materials, trade and reference books, and assessments.

Figure 1 shows the web page of the Houghton Mifflin website where test questions were taken from.
An example of the paper-and-pencil version can be seen in Figure 2.
Figure 2. Paper-and-pencil Version Assessment

A website was constructed to allow students to log in and take online versions of the math tests. The website can be found at http://www.quia.com/pages/tinachan.html. Figure 3 shows the homepage to the website containing all the online tests.
Students would click on the test number and log in using a username and password. Figure 4 shows a sample online test question from Test 10.
Data Analysis

The end result of the research included data analyzed into charts, graphs, and written format. The main statistical tool used in this research study was finding the averages. Standard deviations also played a crucial role in this research. Knowing the standard deviation helped create a more accurate picture of the distribution of means. Furthermore, variables included gender and student’s computer experience. Including these two sets of data will allow for triangulation.

Figure 4. Sample Online Test Question
Summary

This chapter has explained the methods used in this quantitative study of comparing two different math assessment tools for third-graders. The next chapter presents the results obtained with those methods.
CHAPTER FOUR
RESULTS AND DISCUSSION

Introduction

This study was constructed to examine online testing and the effects of different assessment tools on student scores and student learning. Third-grade students were tested to see whether they performed better on computer tests or traditional paper-and-pencil tests. Results revealed similar findings in previous studies.

Presentation of the Findings

Student Math Test Results

Third-grade students completed a total of fifteen math online tests and fifteen math paper-and-pencil tests. Scores for the online tests were automatically scored by the online assessment program, while paper-and-pencil tests were hand-graded. The overall results seem to indicate that students averaged a little better on computer tests than paper-and-pencil tests. The difference in mean is 2% (See Table 4).
Table 4. Overall Mean Scores for Computer-based Test versus Paper-and-pencil Test

<table>
<thead>
<tr>
<th>Assessment Tool</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer-based Test</td>
<td>69.9</td>
<td>8.6</td>
</tr>
<tr>
<td>Paper-and-pencil Test</td>
<td>71.9</td>
<td>8.1</td>
</tr>
</tbody>
</table>

Third-grade students took one computer test and one online test each day, for a total of fifteen online tests and fifteen paper-and-pencil tests. Each test was comprised of 8 multiple choice questions. These math questions were from the student’s own school math curriculum, known as Harcourt Math. Students were taught third-grade math concepts all year round. The math tests were a review of third-grade math standards.

As seen in Figure 5, about half of the daily tests averaged close together, while the other half showed immense differences in averages. Based on these findings, looking at each daily test average is not adequate. Overall averages served a better representation.
In regards to gender, of the seventeen participants 65% were females and 35% were males. Girls did significantly better than boys. On the paper test, girls averaged 73% while boys averaged 64%, with a mean difference of 9%. On the online test, girls averaged 75% while boys averaged 67%, with a difference of 8%. Overall both genders did their best on the online tests (Figure 6).
Student Survey Results

The second part of the research included a survey. This survey was given to the third-graders after they had completed all the online and paper math tests. The survey included five questions directly related to their past computer experience and also their attitudes towards taking the math assessments online and on paper-and-pencil.

All the results of the five survey questions were analyzed in a similar way. Out of seventeen students, ten students averaged best on the paper versions of the test.
and only seven averaged best on the computer versions of the test.

Table 5. Results of Question One

<table>
<thead>
<tr>
<th>Question 1: How often do you use computers?</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) I use a computer often, both at home and at school.</td>
<td>5 (50%)</td>
<td>2 (20%)</td>
<td>3 (30%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>b) I use a computer sometimes, mostly at home.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) I use a computer sometimes, mostly at school.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) I never or almost never use a computer.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When students who scored higher on paper tests were asked how often they used computers, a majority of them (50%) answered that they used computers often, both at home and at school. When students who scored higher on the computer tests were asked the same question, a majority of them (74%) also answered that they used computers often, both at home and at school (Table 5). These outcomes reveal that all the students have some prior basic computer experience.

For the second survey question, students were asked how they felt about using computers (Table 6).
Table 6. Results of Question Two

<table>
<thead>
<tr>
<th>Question 2: How do you feel about using computers?</th>
<th>a) I feel I do my best work on a computer.</th>
<th>b) I am comfortable using a computer.</th>
<th>c) Using computers does not appeal to me.</th>
<th>d) I am frequently frustrated when I use a computer.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>a)</td>
<td>b)</td>
<td>c)</td>
<td>d)</td>
</tr>
<tr>
<td>Students who scored higher on paper test</td>
<td>4(40%)</td>
<td>6(60%)</td>
<td>0(0%)</td>
<td>0(0%)</td>
</tr>
<tr>
<td>Students who scored higher on computer test</td>
<td>5(72%)</td>
<td>2(29%)</td>
<td>0(0%)</td>
<td>0(0%)</td>
</tr>
</tbody>
</table>

For the ten students who scored higher on the paper tests, the majority of them (60%) answered they were just comfortable using a computer and not necessarily doing their best work. Whereas the seven students who scored higher on the computer tests answered that most of them (72%) felt they did their best work on the computer. These results show that students who felt they did their best work on the computer scored higher on computer tests than students who felt they were just comfortable using a computer.

The third survey question asked students how they felt about taking future state tests on the computer (Table 7).
Table 7. Results of Question Three

<table>
<thead>
<tr>
<th>Response</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students who scored higher on paper test</td>
<td>3(30%)</td>
<td>2(20%)</td>
<td>0(0%)</td>
<td>5(50%)</td>
</tr>
<tr>
<td>Students who scored higher on computer test</td>
<td>4(57%)</td>
<td>1(14%)</td>
<td>1(14%)</td>
<td>1(14%)</td>
</tr>
</tbody>
</table>

For the students who performed better on the paper tests, 50% of them answered that it would be difficult to take state tests online at first, but would become eventually become easier once they got used to it. As for the students who scored higher on the computer tests, 57% of them responded that taking state tests on the computer would be easier and more enjoyable than a paper and pencil test. These findings show that students who feel computer tests are easier and more enjoyable tend to do better on computer tests than those who find computer tests difficult at first.

The next survey question asked students how well they thought they did on the computer tests (Table 8).
Table 8. Results of Question Four

Question 4: On the computer part of the test
- a) I did my best work.
- b) I didn't do as well as I did on the paper and pencil part of the test.
- c) I did equally well on the computer and paper and pencil parts of the test.
- d) I was so frustrated using the computer; I was unable to do my best work.

<table>
<thead>
<tr>
<th>Response</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students who scored higher on paper test</td>
<td>3(30%)</td>
<td>0(0%)</td>
<td>7(70%)</td>
<td>0(0%)</td>
</tr>
<tr>
<td>Students who scored higher on computer test</td>
<td>6(86%)</td>
<td>0(0%)</td>
<td>1(14%)</td>
<td>0(0%)</td>
</tr>
</tbody>
</table>

Seventy percent of students scoring higher on the paper tests answered that they thought they did equally well on both computer and paper tests. On the other hand, 86% of the students who scored higher on the computer tests answered that they did their best work on the computer versions of the test. These findings show that students who felt comfortable on the computer and did their best work on the computer ultimately scored higher on the computer tests than the paper tests.

The final survey question dealt with time needed to take the computer tests (Table 9).
Table 9. Results of Question Five

Question 5: The time needed to take the computer part of the test was...
   a) less than a paper and pencil test would have taken.
   b) more than a paper and pencil test.
   c) about the same as a paper and pencil test.
   d) I didn’t notice how long the computer test was taking.

<table>
<thead>
<tr>
<th>Response</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students who scored higher on</td>
<td>5 (50%)</td>
<td>1 (10%)</td>
<td>2 (20%)</td>
<td>2 (20%)</td>
</tr>
<tr>
<td>paper test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students who scored higher on</td>
<td>1 (14%)</td>
<td>2 (29%)</td>
<td>2 (29%)</td>
<td>2 (29%)</td>
</tr>
<tr>
<td>computer test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Of the students who scored higher on the paper tests, 50% of them responded that the computer tests took less time than the paper tests. Interestingly, the students who scored better on the computer tests had mostly different answers to the question on time. Table 9 shows that 29% of them said that computer tests took longer than the paper tests, 29% of them said that the computer test took longer than the paper tests, and 29% said they didn’t notice how long the computer test was. Looking at these results, it is clear that the students who didn’t score so well on the computer tests said it took less time to do computer questions than paper questions. Time is always a factor in how well you do on a test.
These were the results and findings of the research. The following section will give a detailed discussion of each finding.

Discussion of the Findings

The question of equivalence has been asked many times before. The findings in this research study support the findings of other research studies. Research by Clariana (2002) indicated that the freshman business undergraduates who took the computer-based test outperformed those who took the paper-based test. In my research, a similar scenario appeared.

The first set of quantitative results revealed that the overall mean scores for computer-based tests and paper-and-pencil tests were very similar (Table 4). The overall results indicate that students averaged a little better on computer tests than paper-and-pencil tests with a difference in mean of only 2%. Therefore, these results seem to indicate that there is no real significant difference in either assessment tools. Taking a multiple-choice math test on the computer would likely give the same results as taking a multiple-choice math test on regular paper and pencil. These results would also concur with Olsen, Maynes, Slawson, and Ho’s (1986) study.
that showed no significant difference existed between paper-administered tests and computer-administered tests. These findings are very relevant for school districts who decide to go paperless in the future.

Gender is another issue related to education and testing. Research has indicated that American boys are in an academic crisis, while girls have improved their test performance (Mead, 2006). Vanderbilt University were also involved in a study and discovered that females have a significant advantage over males on timed tests and tasks—especially among preteens and teens (2006). My research affirms these findings. Of the seventeen participants, 65% were females and 35% were males. On the paper test, girls averaged 73% while boys averaged 64%, with a mean difference of 9%. On the online test, girls averaged 75% while boys averaged 67%, with a difference of 8%. Overall, girls did significantly better than boys on both versions of the tests.

After the quantitative part of the research, a follow-up survey was given to the third-graders to evaluate and interpret the quantitative results. The survey included five questions directly related to their past computer experience and also their attitudes towards taking tests with different assessment tools. The first
survey question asked students about their past computer experience (Table 5). All students responded as having some computer experience, either at home or at school. The results affirmed McDonald’s 2002 study which argued that the extent to which test takers have experience of using computers ultimately influences their performance on computerized tests. Taylor, Kirsch, Eignor, and Jamieson (1999) found that computer experience encompassed “frequency of use, type of use, number of courses involving computers, owning a computer, access to computers, attitudes towards computers, and related technologies. (pp. 219-274).”

The second survey question asked students how they felt about using computers. The students who scored best on the computer tests all had prior computer experience. Seventy-two percent of these felt they did their best work on a computer. McDonald defines computer anxiety as the fear experienced when interacting with a computer or anticipating an interaction with a computer (2002). The students in my research did not have computer anxiety or lacked familiarity with computers. These findings also coincide with Park’s (2003) study with third-graders. Seventy-nine percent of the third graders surveyed believed they had done their best work on the
computer-based reading test, compared with a little more than five percent of third graders who felt they had done better on paper or were frustrated using the computerized test. Sixty-two percent reported that the online reading test was easier to use and more enjoyable than the paper test, and fifty-eight percent said the same about the Web-based test in mathematics.

The No Child Left Behind legislation has increased educator’s interest in low-to-no-stakes computer-based assessments. These low-to-no-stakes computer-based assessments help administrators and educators prepare students for the high-stakes state exams at the end of the school year. The possibility of taking state online exams will most likely occur in the near future. The next survey question asked students how they would feel about taking a state test on the computer. The majority of students who did well on the paper tests responded that it would be more difficult at first, but would then become easier. Furthermore, the majority of students who scored best on the computer tests responded that taking state tests on the computer would be easier and more enjoyable than a paper test. These findings correlate to the informal survey given to Oregon students. The survey asked Oregon students what they thought about the state’s computerized
assessments. The survey revealed that students found computer-based testing faster and more enjoyable than paper-and-pencil tests. The students also reported feeling more confident on their performance on computerized assessments than on traditional tests (Park, 2003).

The fourth question on the survey asked third-graders if they did their best work on the computer assessments (Table 8). Eighty-six percent of the students who did better on the computer tests responded that they had indeed put their best effort into the computer tests. These findings show that the students who were less frustrated on the computer tests ultimately did their best on that version of the tests.

The last survey question had results that did not entirely match the Oregon students’ responses. My third-graders were asked if the computer tests took longer than the paper tests (Table 9). Fifty percent of the students who scored higher on paper tests responded that it took less time to complete the computer tests than the paper tests, whereas only fourteen percent of the students who scored higher on the computer tests responded that the computer tests took less time. This might seem unusual at first but a deeper look would tell us that the students who took their time on a test ultimately score better.
Therefore, we can conclude that the students who scored higher on the computer tests did well because they thought it took longer to complete the tests.

Summary

One third-grade classroom consisting of seventeen students from Michael D'Arcy Elementary in Fontana, California were assessed. Findings revealed that 59% of the students did better on the paper tests and 41% of the students did better on the computer tests. Further analysis revealed that the overall average of the computer tests was 71.9%, while paper tests revealed an average of 69.9%. In summary, these findings report that there is no significant difference in scores when taking a test on the computer or a test on paper.
CHAPTER FIVE
CONCLUSIONS AND RECOMMENDATIONS

Introduction

The inspiration behind this research stemmed from my own experience in the classroom. I have spent countless hours grading papers and exams. The awareness of online assessments was a light into my dark world. With this new light, I began sharing with my colleges the new light I had found. There is no need to calculate grades any more, the computer does it! There is no need to generate your own progress reports, the computer does it! There is no more need to grade papers by hand; the computer can do it for you!

This research thesis has further made me aware of the endless possibilities that online assessments can have on education. This final chapter gives an overview of the research and its findings. It will also give recommendations for future studies.

Conclusions

Research Questions and Hypothesis Conclusions

The move towards online assessments has raised many questions to whether different assessment tools affect
student scores and student learning. The study began with three specific research questions:

- Will students score higher on the computer test than the paper-and-pencil version?
- If students were given the choice, would they prefer to take tests on the computer rather than the more traditional medium?
- Does gender play a role in how well students succeed on the computer tests?

Using convenient sampling, one third-grade classroom consisting of seventeen students from Michael D’Arcy Elementary in Fontana, California were assessed. The research activities covered a three-week period, from June 5, 2006 to June 23, 2006. During this time, these third-grade students served as their own control group by taking several math tests online and several math tests with paper and pencil. Results were then compared to see if performance on computer-based tests would be more successful than pencil-and-paper tests. The overall results seem to indicate that students averaged a little better on computer tests than paper-and-pencil tests. The difference in mean is 2%, thus revealing that there is no significant difference in computer test scores and paper
test scores. These findings were consistent with other research studies.

The second part of the research included a survey. This survey was given to the third-graders after they had completed the math tests. The survey included five questions directly related to their past computer experience and also their attitudes towards taking the math assessments online and on paper-and-pencil. The third survey question in particular, asked students how they felt about taking future state tests on the computer. For the students who performed better on the paper tests, 50% of them answered that it would be difficult to take state tests online at first, but would become eventually become easier once they got used to it. As for the students who scored higher on the paper tests, 57% of them responded that taking state tests on the computer would be easier and more enjoyable than a paper and pencil test. These findings correlate to the informal survey given to Oregon students. The survey asked Oregon students what they thought about the state's computerized assessments. The survey showed students finding computer-based testing faster and more enjoyable than the traditional paper and pencil tests. The students also reported feeling more
confident on their performance on computerized assessments than on traditional tests (Park, 2003).

The study permitted examination of the two following hypotheses:

• There will not be a difference in scores between students who take the math tests on the computer and those who take it on paper.

• If given the choice, students would not prefer to take tests on the computer rather than the more traditional medium.

• The difference in gender does not play a role in how well students succeed on computerized tests.

The first hypothesis was mostly correct. As mentioned before there was only a 2% difference in mean scores between the computer test and the paper test. Therefore we can say there is no significant difference between students’ scores on computer and paper tests. The second hypothesis was not a correct assumption. A majority of the third-graders either answered that taking state tests on the computer would be difficult at first but would eventually become easier once they got used to it or that the computer tests were ultimately easier and more enjoyable than the paper tests. Lastly, the third hypothesis was a false assumption. In regards to gender,
of the seventeen participants, 65% were females and 35% were males. Girls did significantly better than boys. On the paper test, girls averaged 73% while boys averaged 64%, with a mean difference of 9%. On the online test, girls averaged 75% while boys averaged 67%, with a difference of 8%.

Meaning of Study

The results of these findings are very crucial to the future of education and online assessments. Because of the increased testing requirements for the No Child Left Behind Act, school districts will be forced to find more alternative ways to assess students. Teachers will use low-to-no-risk computerized assessments to improve instruction and provide remediation for students who are failing. The increased testing requirements for the No Child Left Behind Act will completely revolutionize online assessments in the 21st century. Districts will jump on the bandwagon of computerized assessments and as a result instruction will improve and remediation will exist for at-risk students.

There is no question that testing by computer has arrived. Since 1994, nursing licensure examinations have been given solely on computers. A large number of organizations and associations that grant certification

60
and licensure are either currently giving tests by computer or are seriously considering doing so (Bugbee, 1996).

Given current trends in education and technology, the timeliness and relevance of this research cannot be overlooked. The educational community needs to know whether computer-based testing is equivalent to paper-and-pencil testing.

Future Research and Recommendations

The findings based on this research and comparable literature reviews leave a hole in certain areas of online assessments. Suggestions for further research and study are made to further the area of computer-based assessments.

Computer adaptive testing is a relatively new area of study. Computer Adaptive Tests, or CATs, are quite different from the computer standardized tests seen in the past. The main difference between CATs and computer tests is that CATs “adapt” to the performance of the students. Each test taker is given a different mix of questions depending on how well he or she is doing on the test. This means the questions get harder or easier depending on whether questions are answered correctly or not. Score is
not only determined by how many questions were answered right, but by the difficulty level of these questions. Further study in this area will enhance the future of computer-based testing.

Another area of study would include computer testing time. This study would question the speed it would take to complete reading comprehension tests and compare scores to take reading comprehension tests on paper. Another focus would be figuring out which test took longer to complete. Time students to see how long they take to complete the computer tests and the paper tests. These timed scores can give valuable insight into whether one test was finished sooner than the other.

Gender difference and the role it plays in education has been a current issue in the field of online-assessments. Why do girls perform better than boys in school? Most recently, why do girls perform better than boys in timed tests? This would be another possible future research study.

Finally, computer experience would be another interesting topic to delve into. Instead of observing third-graders, high school students might give a better insight into how computer experiences reflect scores on computer-based assessments.
Summary

The No Child Left Behind Act has driven interest in computer-based assessment products. Schools now use assessments as an ongoing method of monitoring progress and linking data to instruction. Some observers expect use of computer-based testing to rise steeply over the next few years. With the interest in computer-based assessments rising, it is necessary to reach a full understanding of the equivalence of computer tests and paper tests. Once this is accomplished the future of instructional technology seems endless.
APPENDIX A

PARENT INFORMED CONSENT
A COMPARISON STUDY OF MATH ASSESSMENT TOOLS FOR THIRD-GRADERS

PARENT INFORMED CONSENT

The study in which your child is being asked to participate was designed to investigate whether or not different assessment tools affect student learning and student achievement. This study will seek to answer two questions:

1. Will students score higher on computer-based math tests than traditional paper-and-pencil tests?
2. If given the choice, will students prefer to take tests on the computer rather than on the more traditional medium?

This study is being conducted by Mrs. Tina Chan under the supervision of Eun-Ok Baeck, PhD, professor of the College of Education, Department of Instructional Technology. This study has been approved by the Institutional Review Board, California State University, San Bernardino.

In this study your child will be asked to take 15 math tests on the computer and 15 math tests on paper. There will be 8 questions on each test for a total of 240 questions. The task should take about 15 minutes per day for approximately one month. All of your child's responses will be held in the strictest of confidence by the researcher. Their name will be reported with their responses online at http://www.quip.com/paperdirection.html. Only the researcher, parents and students can view scores with a valid username and password. Parents can view all copies of tests and information about the online program on the website.

No potential risks will occur from the procedures used in this study. On the other hand, teacher, students, and parents will benefit greatly from this study. Teachers, students, and parents will get immediate results from the tests and will be able to use this valuable resource to pinpoint student's strengths and weaknesses in math. Student's math skills will also improve.

Your participation in this study is totally voluntary. You are free to withdraw at any time during this study without penalty. In order to ensure the validity of the study, we ask you not to discuss this study with other students.

If you have any questions or concerns about this study or would like to receive the results of the study, please feel free to contact Mrs. Tina Chan or Professor Eun-Ok Baeck, PhD at (909) 537-5454.

By placing a check mark in the box below, I acknowledge that I have been informed of, and that I understand, the nature and purpose of this study, and I freely give consent to my minor child to participate. (ECF 51513 requires that the parent or guardian of the pupil be notified in writing that this test, questionnaire, survey, or examination is to be administered and that written approval by parent or guardian gives pupil permission to take test, questionnaire, survey, or examination.)

Place a check mark here □

Student Name: ____________________________

Today's Date: ____________________________

Parent/Guardian Name: ____________________________

Signature: ____________________________

Student Name: ____________________________

Today's Date: ____________________________

Parent/Guardian Name: ____________________________

Signature: ____________________________
APPENDIX B

STUDENT ASSENT FORM
A COMPARISON STUDY OF MATH ASSESSMENT TOOLS FOR THIRD-GRADE STUDENTS

STUDENT ASSENT

You are being asked to participate in a study that will test to see if students will score higher on a computer test than on a regular paper-and-pencil test.

Mrs. Tina Chan is doing this study with the help of Eun-Ok Back, PhD, professor of the College of Education, Department of Instructional Technology. This study has been approved by the Institutional Review Board, California State University, San Bernardino.

In this study you will be asked to take 15 math tests on the computer and 15 math tests on paper. There will be 8 questions on each test for a total of 240 questions. All your answers will be held private.

No potential risks will occur from the procedures used in this study, but you will improve your math skills.

Your participation in this study is totally voluntary. You are free to withdraw at any time during this study without penalty. In order to ensure the validity of the study, we ask you not discuss this with other students.

If you have any questions or concerns about this study or would like to receive the results of the study, please feel free to contact Tina Chan at (909) 649-6908 or Professor Eun-Ok Back, PhD at (909) 880-5454.
REFERENCES


