The Relationship of Testing Mode and Mathematical Assessment Performance

By Stephanie Layden

Submitted in Partial Fulfillment of the Requirements for the

Degree of Master of Education

May 2018

Graduate Programs in Education

Goucher College

Table of Contents

List of Tables

Abstract

The purpose of this quasi-experimental study was to examine the impact of test format on student assessment performance. The study focused on the Mathematics District Assessments 1 and 2 after controlling paper-and-pencil platform in Assessment 1 and receiving either a technology-enhanced or paper-and-pencil format in Assessment 2. The null hypothesis was that the assessment performance of the 3rd grade students who were assessed using a technology-enhanced assessment would not be significantly different from the performance of those who were assessed using the paper-and-pencil assessment. The null hypothesis was rejected, as the assessment performance of elementary school students who were assessed using a paper-and-pencil assessment was significantly better than the performance of those who were assessed using the technology-enhanced assessment.

CHAPTER I

INTRODUCTION

Children growing up in today's society are faced with the constant shifting of technology trends, some of which will have a positive consequence and others negative. From the traditional classroom filled with text books, literature libraries, and paper-and-pencil assessment, classrooms are moving to the technology advanced classroom with SMART boards, Chromebooks, and internet access. In a survey conducted by the Bill and Melinda Gates Foundation (2014), teachers reported that 39% of the time technology played the primary role in delivery of mathematics instruction with an additional 32% of the time being spent using technology to support instruction.

The integration of technology into classroom instruction has been an important issue for many years. Today's students are more diverse than ever, and technology inherently allows teachers to individualize instruction and target specific content or skills. As stated in the Edutopia video (2012), "An Introduction to Technology Integration," integrating technology with classroom practice can be a great way to "strengthen engagement by linking students to a global audience, turning them into creators of digital media, and helping them practice collaboration skills that will prepare them for the future".

Classroom instruction and assessment go hand in hand. With the infusion of technology into classroom instruction, it is also important to rethink the way assessments are delivered. In 2009, only 32 percent of middle school students said that they were taking tests online (Project Tomorrow, 2014). Teacher facilitation of online tests in middle school classrooms increased to 47 percent by 2013, "a strong statement about both the availability of online tests within curriculum and teacher adoption of those tools for assessment purposes" (p.3). This technology

shift involves the teacher-focused classroom paradigm transforming into a student-centered, technology-enhanced project-based learning environment to develop critical thinking, problem-solving and teamwork skills. This will prepare students for college, career and the workplace.

The researcher's interest in this topic was a direct result of her current position as an Elementary Math Specialist for Anne Arundel County Public Schools. The researcher works with grades k-5, however grade 3 is the researcher's "grade of focus." Last year, schools were permitted to decide if 3rd grade students would take the PARCC assessment online or paper-and-pencil. This made the researcher question if one of these modes would be of greater benefit for student achievement. Currently, schools in Anne Arundel County are able to choose to do the County District Assessments online or by paper-and-pencil. Because assessment data is very relevant to the researcher's job, she was curious about the impact of this decision.

Statement of Problem

The purpose of this study is to examine the impact of technology-enhanced assessments on mathematical testing data of grade 3 students in Anne Arundel County Public Schools on the Elementary Mathematics District 2 Assessments.

Hypothesis

The null hypothesis was that the assessment performance of elementary school students, who were assessed using a technology-enhanced assessment, would not be significantly different from the performance of those who were assessed using the paper-and-pencil assessment.

Operational Definitions

Paper-and-Pencil Assessment is defined as an assessment that students take on paper. There is no part of the assessment that involves the use of technology. Students complete the assessment using the test paper and a pencil.

Technology-enhanced is defined as an assessment that is completed using a computer or tablet. Students may have scratch paper, but only what is submitted on the computer or tablet will be assessed.

Assessment Performance is defined as the comparison between the score of the first quarterly District Assessment with the score of the second quarterly District Assessment.

CHAPTER II

REVIEW OF LITERATURE

This literature review explores the impact of technology-based assessments on mathematical assessment data. Section one presents a brief overview of technology and its relationship to curriculum and instruction. Section two discusses technology in the mathematics classroom. Section three provides a definition and overview of technology-based assessment. Section four examines the connections between the platform of an assessment and the student's performance.

Technology, Curriculum and Instruction

Curriculum and instruction are ideally designed to use best practices to prepare quality standards-based content, learning experiences, and assessments to evaluate students' preparation for the 21st century workforce. The use of technology to support curriculum and instruction has steadily increased. As Moore (2003) noted, technology has provided a tool "to link academic standards with higher order thinking skills that are necessary in properly assessing students' mastery of standards" (p. 22). Mastery of standards can be assessed on an individual basis or on the conceptual understanding that a student demonstrates when working with a group of students. The value of technology in the hands of students is clear. Yuan-Hsuan, Waxman, Jiun-Yu, Michko, and Lin (2013) noted that working with technology individually offers greater flexibility by allowing individuals to work at their own pace. Research has shown that students working in small groups (e.g. 3-5) with computers performed better than an individual student working with computers. In a survey on the availability and use of educational technology among teachers in public elementary and secondary schools in the U.S., Gray, Thomas, and Lewis (2010) found that the ratio of students to computers in the classroom was 5.3 to 1. Though

computers allows students to work at their own pace, which has benefit, collaboratively working with peers has even greater benefit. Whether working individually or in small groups, students should have access to technology to deepen their conceptual understanding of the academic standards (Moore, 2003).

In order to prepare students for the future and help them learn to think, learn, and gain different perspectives, technology has to be integrated into the classroom (Eyyam & Yaratan, 2014). In the State of Maryland, students are assessed each year using a high stakes technology-based assessment. The Partnership for Assessment of Readiness for College and Careers (PARCC) is a group of states working together to develop a set of assessments that measure whether students are on track to be successful in college and careers. In order to prepare students for high stakes mandated testing that is technology-based, students must be exposed to technology-based assessment. When students encounter the PARCC Assessment, they will be expected to answer questions by using the drag and drop function, to highlight important information as they read, to watch videos, to fill in boxes in a table, as well as to type in responses. There are many different mouse-manipulation tasks, each requiring a degree of comfort with technology. Becoming comfortable with technology takes time and practice, and practice with online assessments is crucial. Hain (2011) noted that high quality assessments given throughout the year help educators determine where and when extra support and enrichment are appropriate. Moore (2003) explained that teachers are responsible for producing knowledgeable and competent students, and when teachers are involved in the changing of assessment, they tend to be more detailed in planning. Detailed planning in turn produces higher quality instruction and assessment. Technology integration is changing the traditional learning environment, which in turn impacts the way students should be assessed.

Technology in the Mathematics Classroom

Research has suggested that when technology is integrated into instruction, both teachers and students benefit (Eyyam & Yaratan, 2014). In order to ensure 21st century readiness for all students, schools must teach 21st century skills. The Partnership for 21st Century Learning developed a Framework for 21st Century Learning with input from teachers, experts in education, and business leaders to define the skills and knowledge students need to succeed as citizens and workers in today's digitally interconnected society. These skills include content knowledge, learning and innovation skills, life and career skills, and information, media, and technology skill (Partnership for 21st Century Learning, 2007).

Mathematics is a fundamental subject that is essential for students in the 21st century. Additionally, teachers are tasked with building deep conceptual understanding rather than shallow knowledge. Teachers use Chromebooks, laptops, SMART boards, document cameras, and other technologies to bring mathematics instruction to life and create meaningful learning experiences for students. Providing meaningful integration of new technologies integrated with the best instructional practices provides a powerful learning experience (Holland & Holland, 2014). Eyyam and Yaratan (2014) noted that in order to prepare students for the 21st century future and encourage them to think, learn, and appreciate different perspectives, technology must be integrated into classroom instruction. Some examples of this could be the use of SMART boards during large or small group instructions, to allow all pupils to respond to questions, and Chromebooks to enable students to independently practice skills. With technology available, teachers have increased ability to provide varied instruction and access to differentiated curricular materials to meet the needs of diverse learners in the classroom (Holland & Holland, 2014).

Nguyen, Hsieh, and Allen (2006) cite several previous studies that have noted that students who use computer-based learning found mathematics more enjoyable. When interviewed about their experiences with computer-supported mathematics instruction, some students noted that computer math made the problem-solving process easier, gave them more information or clues to use in making sense of the mathematics, that it was more interesting than paper-and-pencil mathematics, and that they felt computers made learning more fun.

Overview of Technology-Based Assessments

In an increasingly complex and information rich society, students and teachers must use technology effectively to live, learn, and work successfully (Eyyam & Yaratan, 2014). Holland and Holland (2014) noted that as a society, we are seeing many new technologies arrive and old ones go away. Reading the newspaper has been replaced by reading the internet. Handwritten letters have been replaced by typed emails or texts. But why are these changes so relevant and even crucial to educational needs and reform? In 2001, the Elementary and Secondary Education Act was reauthorized by U.S. Congress and became known as The No Child Left Behind Act (NCLB). According to NCLB (Dee & Thomas, 2010), states must develop basic skills assessments for annual testing in grades 3-8 in order to receive federal funding. As noted by Poggio, Glasnapp, Yang, and Poggio (2005), NCLB designated assessment as a cornerstone of gauging school success. Moore (2003) stated that with the infusion of technology in schools, the use of technology allows students to be assessed in ways that will measure "their ability for connecting knowledge learned with real-world applications" (p. 22).

With funding and success hinging on their results, the topics of assessment and assessment data are on the minds of students, teachers, parents, and administrators alike. Assessment can be considered a systematic process of making judgements and reporting results

about effectiveness of learning or an individual students' progress toward mastery of educational objectives (Laborda, Sampson, Hambleton, & Guzman, 2015).

One way to assess students is through the use of technology. "Technology or computer-based assessment is a term used when a computer delivers a test in the same manner as it would appear if it were in the form of a paper-and-pencil test" (Maguire, Smith, Brallier, & Palm, 2010, p. 117). Technology-based assessments (TBA) have many benefits to teachers and districts. Poggio et al. (2005) noted that some benefits of TBAs include immediate score reporting on student performance, the reduction in costs related to printing, shipping, and administering paper-and-pencil (P & P) assessments, test security improvements, as well as the opportunity for continuous testing of students. TBAs allow teachers to administer assessments while ensuring that all students are given the allocated time for completion. TBAs can also easily allow for students to have or be denied access to support tools, such as calculators or text-to-speech tools as appropriate.

The Partnership for 21st Century Learning (2007) describes the required balance of technology-enhanced, formative and summative assessments that measure student mastery of 21st century skills. Some examples of formative assessments that incorporate technology might be posting questions on Google classroom, making an exit ticket on Kahoot and having the students use technology to respond, or creating a survey or multiple question exit ticket using Plickers and scanning students' answer cards to graph the data for discussions. In addition, teachers' use of technology can enhance assessments administered for summative purposes. These kinds of assessments could have selected response, drag and drop, equation builder, drop down menus, constructed response, and multiple part questions. Each of these modalities can be incorporated into formative and summative assessments to prepare students for the state

assessments, including PARCC, that they take each year. A major benefit of technology-based formative and summative assessment is that they provide teachers instantaneous data that can be used to clarify misconceptions and guide and differentiate instruction to maximize learning.

Assessment Modes and Performance

Curriculum and instruction also include the element of assessment. Despite gains in the use of technology in instruction and everyday life, Clarke-Midura and Dede (2010) noted "the field of assessment has not progressed much beyond paper-and-pencil assessments" (p.309). They also stated that paper-and-pencil assessments fall short of providing measures of the knowledge and skills students need to be prepared for the 21st century workforce. Paper-and-pencil assessments are often formatted as multiple-choice and short answer questions, which do not yield responses that reflect the higher order type of thinking necessary to demonstrate the competencies necessary for college and career readiness. As a result, "the curriculum is full of low-level facts and recipe-like procedures" (p. 312), limiting the validity of its assessments to reflect students' ability to solve real life problems.

Clarke-Midura and Dede (2010) noted that The National Research Council Report, "Knowing What Students Know," identified three key components to assessment. These are cognition, which refers to a model of the thinking and learning which students demonstrate within a domain; observation, the tasks students engage in that demonstrate evidence of learning; and interpretation, the process for making sense of the evidence. The report stated that when these three components work simultaneously, assessments are effective. It was also noted that paper-and-pencil assessments, which reflect the observation part of the triad typically do not generate rich observations because students often choose among pre-determined options on

them, which does not accurately demonstrate their mastery of a skill involving advanced knowledge.

With advances in technology come advancements in the possibilities for assessment. Technology has provided tools that connect academic standards with the higher-level thinking skills that are necessary to assess learners' competencies. The 21st century is no longer a "penmanship" century in the United States (Nguyen et al., 2006). In a study of the impact of web-based assessment on students' mathematical learning attitudes, it was found that students felt smarter while doing mathematics on the computer and gained more self-confidence when receiving immediate feedback. These responses reached far beyond what paper-and-pencil could do for students.

Current Research

A 2008 study completed by Bennett, Braswell, Oranje, Sandene, Kaplan, and Yan addressed three main areas related to the comparability of computer and paper mathematics test scores for American eighth graders. In regards to performance, the mean score for eighth graders who took the computer test was significantly lower than the mean for students who took the paper version of the same assessment. The second part of the study looked at the performance of selected groups based upon race and ethnicity, parents' education level, region of the country, school location, and school type. The results showed there was generally no difference in the mean performance on the paper versus computer test. The last question of the study dealt with the computer familiarity of students. It was found that scores of students taking the test on the computer were directly related to their familiarity with the computer. As their keyboarding skills increased, their scores increased. The authors did note several limitations in the study. The most important is that the data was collected in 2001, which, when looking at technology, is a

significant amount of time before publication. In addition, when looking at the results for the selected groups, the groups were often small. When the study was designed, the impact of familiarity was assessed for the overall study sample, not the selected groups. Because of these limitations, the authors recommended further research to clarify the relevance of these findings for current students.

In 2014, Eyyam and Yaratan conducted a study in a private school in Cyprus to investigate whether the use of technology in the classroom impacted achievement and attitudes of seventh grade students in mathematics class. The students were broken into five groups, three experimental groups and two control groups. All of the students were given a pre-test and a post-test. The students in the experimental group received instruction using technological tools, and the control group was taught using traditional methods. After the post-test was administered and evaluated, it was noted that students who were instructed using technology had significantly higher test score than those instructed without technology. It was also noted that students preferred to be in the class where technology was used.

In a study of college students at a mid-sized state-supported university in the southeast, researchers examined the difference in test scores for students who took all assessments electronically compared to students who took all assessments through a paper-and-pencil format. In this study, all students received the same instruction; the only difference was the assessment format. Maguire, Smith, Brallier and Palm (2010) noted that students who completed all assessments electronically scored significantly higher than those students who completed all assessments using the format of paper-and-pencil.

In a study conducted by Anne Arundel County Public Schools Instructional Data Division (2016), Grade 3 PARCC assessment data was analyzed in relation to the testing mode:

paper-and-pencil or online. In the spring of 2016, grade 3 students were administered the PARCC assessment. The choice of the assessment mode, paper-and-pencil or online, was left to the administration in each elementary school. When comparing student performance on the PARCC mathematics assessment sub-claims, more students who took the paper format met or exceeded expectations than those who took the online format. The researchers concluded that there is a difference in student performance based upon test format. The results of the PARCC 2016 analysis indicate that it may be beneficial to assess students in grade 3 mathematics using paper format rather than an online format.

Summary

This literature review examined some of the ideas presented in the existing research regarding methods of assessment and how they impact results and instruction. Given the fact that technology and assessment requirements are fluid and evolving rapidly, it is timely for researchers to learn more about these topics to implement accurate and cost-effective assessment strategies.

CHAPTER III

METHODS

A quasi-experimental design was used for the comparability study of student assessment performance. The study focused on the Mathematics District Assessments 1 and 2 after controlling paper-and-pencil platform in Assessment 1 and receiving either a technology-enhanced or paper-and-pencil format in Assessment 2. The independent variable is the platform of the District Assessment 2, which was given either by paper-and-pencil or technology-enhanced. The dependent variable was the assessment performance, which was defined as the comparison between the score of the first quarterly assessment and the score of the second quarterly assessment.

Participants

The participants in this study were 5,193 third grade students from 67 schools at a public school district in the State of Maryland. The school system is the 5th largest in Maryland and the 46th largest in the United States, with over 80,000 students and 5,000 teachers supporting a comprehensive curriculum from Pre-K through 12th grade. The control group consisted of 3,738 third grade students from 48 schools who were administered both District Assessment 1 and 2 with paper-and-pencil modality. The focal group consisted of 1,455 students from 19 schools within the same district were administered the assessment using paper-and-pencil for the first quarterly assessment and technology-enhanced assessment for the second quarterly assessment.

Instrument

Data were collected from school district-provided District 1 and 2 Assessments. District Assessments are mandatory for all third grade students. Each District Assessment is a quarterly-based assessment used to measure a student's proficiency with the Common Core State

Standards that were instructed throughout that quarter. Once District Assessments have been administered, the bubble sheets are scanned into the Performance Matters Database. The Performance Matters Database is used for creating, scoring, and storing common assessments given to students within the district. Performance Matters allows teachers and administrators to have access to data almost immediately, allowing for interpretation and reflection.

Procedure

Students in third grade receive instruction based upon the Common Core State Standards of Mathematics and district-provided curriculum. The district expectation is for all teachers to have similar expectations for all assessments so that the scores are valid and the data has meaning. The researcher ensured that all teachers were provided with assessment administration guidelines to ensure consistency with the testing administration. The guidelines state that assessments are to be given during a three-day window, with a 45-minute time restraint, and scanned within 3 days of administration. Students are permitted to use all mathematical tools that were used during instruction. The permitted tools, also referred to as math manipulatives, are listed in the curricular document. Each third grade student is assessed either with the paper-and-pencil or computer-based assessment.

Students took the mandatory District Assessment provided by their school district in the first and second quarter of third grade for this study. This test was taken in the student's classroom independently. The District Assessment required students to answer Task Type 2 reasoning questions, Task Type 3 modeling questions, as well as multiple-choice questions. The test was scored by third grade teachers using a district-developed rubric to score the written portions and a computer scanner to scan assessments into the Performance Matters Database the

district requires. Once the scores were scanned into the district database, they were analyzed by the researcher.

After eight weeks, all students were given the second District Assessment that was based on the instruction from the second quarter. The assessment was the same format as the first District Assessment and was also developed by the district. The same scoring technique was used. The third grade teachers used a district-developed rubric to score the written portions and scanned the remaining items for scoring. The data was analyzed by the researcher.

Student score changes from the first to the second District Assessment were used to capture the impact of technology-enhanced assessment on student test scores.

CHAPTER IV

RESULTS

The purpose of this study is to examine the impact of technology-enhanced assessments on mathematical testing data of grade 3 students in Anne Arundel County Public Schools on the Elementary Mathematics District 2 Assessments. The researcher sought to examine if students' mathematical assessment data on the second quarterly District Assessment was impacted by the platform of the assessment, technology-enhanced or paper-and-pencil. The results of the assessments were averaged to determine if there was a significant difference in the average score change from District Assessment 1 to 2 between students that took both assessments in the paper-and-pencil platform and those who took District Assessment 1 in the paper-and-pencil platform and District Assessment 2 in the technology-enhanced platform. The results are presented in Table 1.

Table 1

Significance of Modality in Quarterly District Assessment 1 and 2 Data

Group	Test Form	N	Ave Change from DA 1 to DA 2	Mean Difference	Sig. (2 tail)
All	both pp	3738	4.51	3.320	0.000
	pp to online	1455	1.19		Yes (pp > online)

$P \leq 0.05$ indicates strong evidence that there is a statistically significant difference in score change between students who took both assessments in the paper-and-pencil platform and students who switched to online format in DA 2. Students taking the paper-and-pencil assessment outperformed the students taking the technology-enhanced assessment.

The data from this study was further analyzed using the t test for different student subgroups. The t test was used to analyze whether the platform of the assessment, paper-and-

pencil or technology-enhanced, had a statistically significant impact on students of different ethnicities or students receiving special services. The results are presented in Tables 2 and 3.

Table 2

Changes in District Assessment Score from Quarter 1 to Quarter 2 by Ethnicity

Group	Test Form	N	Ave Change from DA 1 to DA 2	Mean Difference	Significance
Asian	both pp	142	4.90	-0.147	0.963
	pp to online	36	5.05		
AF AM	both pp	849	4.67	2.378	0.181
	pp to online	160	2.29		
Hispanic	both pp	641	4.28	1.433	0.366
	pp to online	201	2.85		
Multi-Racial	both pp	268	1.78	2.095	0.461
	pp to online	73	-0.31		
White	both pp	1822	4.86	4.218	0.000
	pp to online	982	0.64		Yes (pp > online)

P ≤ 0.05 indicates strong evidence that there is statistically significant difference in score change when comparing white students who took both assessments in the paper-and-pencil platform with white students who took DA 2 online. Although not statistically significant, paper-and-pencil seems working better for African American, Hispanic, and Multi-Racial groups, but not for Asian.

Table 3

Changes in District Assessment Score from Quarter 1 to Quarter 2 for Special Service Groups

Group	Test Form	N	Ave Change from DA 1 to DA 2	Mean Difference	Significance
FARMS	both pp	1561	3.55	1.894	0.145
	pp to online	315	1.66		
SPED	both pp	312	2.06	2.926	0.290
	pp to online	84	-0.86		
ELL	both pp	484	2.51	-0.786	0.685
	pp to online	132	3.29		

Although not a statistically significant difference, paper-and-pencil seems to work better for Free and Reduced Meals (FARMS) and Special Education (SPED) groups, but not for English Language Learners (ELL). ELL students performed better when receiving the technology-enhanced assessment.

CHAPTER V

DISCUSSION

The purpose of this study is to examine the impact of technology-enhanced assessments on mathematical testing data of grade 3 students in Anne Arundel County Public Schools on the Elementary Mathematics District 2 Assessments. Each student took the District Assessment 1 in the paper-and-pencil platform, while District Assessment 2 was taken in either the paper-and-pencil platform or the technology-enhanced platform. The score change from DA 1 to DA 2 was then compared.

The results provided statistically significant evidence, with p value of 0.000, that overall students who took both assessments in the paper-and-pencil platform outperformed the students who transitioned to the technology-enhanced platform. When the results were looked at more closely through the lens of ethnicity, there was a statistically significant difference at p=0.000 in score change when comparing white students who took both assessments in the paper-and-pencil platform and white students who took online DA 2. The p values for other subgroups (African American: 0.181, Hispanic: 0.366, and Multi-Racial: 0.461) did not show a statistically significant difference, though paper-and-pencil seems to produce a slightly higher score. Though it is not statistically significant, it was noted that Asian students performed better when transitioning to the technology-enhanced platform over the paper-and-pencil platform with a mean difference of -0.147. When looking within the subgroups and further grouping into the category of students receiving special service support, there was not a statistically significant difference in the score change with the transition to the paper-and-pencil or technology-enhanced platform.

According to this data, the null hypothesis that the assessment performance of elementary

students who were assessed using a technology-enhanced assessment would not be significantly different from the performance of those who were assessed using the paper and pencil assessment was rejected.

Implications of Results

When looking at the relationship between the platform of the assessment and the assessment data, it is important to look at how this relationship impacts the students. After analyzing the results of this study, there was statistically significant evidence to support the use of paper-and-pencil assessments when looking at the group as a whole. When looking across the subgroups, specifically looking at ethnicity and groups receiving special services, African Americans, Hispanics, Multi-Racial, Whites, FARMS, and Special Education students all had higher average growth with the continued use of the paper-and-pencil assessment format. Two subgroups, Asians and ELL students, showed greater average growth when transitioning from the paper-and-pencil assessment to the technology-enhanced assessment. This data is very important to consider as a teacher. When thinking about the best interests of students and the use of assessment data to drive instructions, it is necessary to consider the platform of the assessments that are placed in front of students.

In addition, there may be implications for practice with technology-enhanced practice assessment items to practice using the technology tools along with the technology platform. State testing in Maryland is currently technology-enhanced. Because of this, it is important for students to have a level of comfort with the use of technology as well as the tools necessary to answer the different types of questions. Practice built into the instructional day would allow students develop confidence with the technology-enhanced platform.

Threats to Validity

Throughout this study, the focus was the impact of the use of paper-and-pencil or technology-enhanced assessments on mathematical assessment data. In analyzing the results of the study, there were a few threats to the validity that were noted. The threats to validity for this study include proficiency with technology use, the ability to show the work associated with solving the problem, the consistency of question format between the two assessment platforms, and the time allotment for the assessment.

Proficiency with technology is a threat to validity because some students may be more fluid users of technology than others. The assessments are intended to provide data on a students' mathematical content knowledge, not their ability to use technology. If students have a technology accessible in school and at home, they will be more comfortable than students that only have access to technology in school. Another factor that could affect validity is the ability to show the work or the process used to show the problem. When an assessment is paper-and-pencil, the student can simply write an equation or draw a model to solve the problem in the test book. When that same problem is on a computer, the student has to self-select to use scratch paper to show work. Otherwise, there is a drawing tool that can be used on the computer.

In addition, the consistency of the question format and response mode between the two assessments, paper-and-pencil and technology-enhanced, is a threat to validity. For example, on the District Assessment 2, students had to use the drawing tool five times to show their work. One example was constructing a line plot. On the technology-enhanced assessment, students had to use their finger to draw each x on the line plot. Students had to plot eight different lengths, each length requiring an x on the line plot. Four other questions asked students to show (draw) or explain (type) their thinking in the box. The demands of these items relate to the last threat of

validity, time. According to the assessment protocol, students are only permitted forty-five minutes to complete the assessment. The demands of the online assessment and the differences in the response mode could cause a student not to complete the assessment, which would lead to a lower score.

Connections to Previous Studies/Existing Literature

In a similar study conducted by Anne Arundel County Public Schools Instructional Data Division (2016), grade 3 PARCC assessment data was analyzed in relation to the testing mode: paper-and-pencil or online. When comparing student performance on the PARCC mathematics assessment sub-claims, more students who took the paper format met or exceeded expectations than those who took the online format. The researchers in this study formed similar conclusions to the current study, there is a difference in student performance based upon test format and it may be beneficial to assess students in grade 3 mathematics using paper format rather than an online format.

A 2008 study completed by Bennett et al. had results that were relevant to the current study. The previous study noted that the mean score for eighth graders who took the computer test was significantly lower than the mean for students who took the paper version of the same assessment. This is comparable to the current study when looking at the group as a whole. In the current study the mean average of students that took both assessments using the paper-and-pencil format was higher when compared to students who transitioned from paper-and-pencil to the online format. The previous study also dealt with the computer familiarity of students. It was found that scores of students taking the test on the computer were directly related to their familiarity with the computer. As their keyboarding skills increased, their scores increased. This directly relates to the current study in regards to the threats for validity. The current researcher

noted that students' familiarity with technology could potentially impact the score of students on the assessment.

In contrast, a study of college students, researchers examined the difference in test scores for students who took all assessments electronically compared to students who took all assessments through a paper-and-pencil format. In this study, all students received the same instruction; the only difference was the assessment format. Maguire et al. (2010) noted that students who completed all assessments electronically scored significantly higher than those students who completed all assessments using the format of paper-and-pencil. This study did look at the same relationship, the format of the assessment compared to the test score. The major differences are the age of students, the fact that all students received the same instruction, and that all assessments were taken in the same format: paper-and-pencil or online. In the current study, there is no way to ensure that all students received the same instruction, though all teachers have access to the same curriculum. In addition, in the current study, all students took the paper-and-pencil assessment for the first District Assessment and then either took the second District Assessment paper-and-pencil or transitioned to technology-enhanced.

Implications for Future Research

Completing this study and analyzing the results has led to a number of new questions that could be explored in future research. With the push to move towards technology-enhanced assessment, it would be interesting to look at the performance of students on the technology-enhanced state assessment compared to their performance on technology-enhanced district assessments. In addition, it would also be advantageous to look at the impact of technology infusion within instruction and the impact that it has on assessment data. Specifically, if a teacher incorporates technology into daily practice, what impact does it have on the performance of

students on technology-enhanced assessments? Lastly, it would be important to look at the students' comfort level with technology and how that correlates with assessment performance.

Conclusion/Summary

This study targets a question that confronts mathematics educators, administrators, and policy makers in the 21st century: What is the impact of technology-enhanced assessments on mathematical testing data? According to this study, there is a statistically significant difference when comparing the average change in data of the two assessment platforms. With that in mind, it is essential to continue exploring the connections between technology use and the impact on students' assessment data.

References

AACPS Instructional Data Division (2016). *Grade 3 PARCC math: Online v. paper format*, Retrieved from https://intranet.aacps.org/Dept/AAR/Research/Pages/Default.aspx

Bennett, R.E., Braswell, J. Oranje, A., Sandene, B., Kaplan, B. & Yan, F. (2008). Does it matter if I take my mathematics test on computer? A second empirical study of mode affects in naep. *The Journal of Technology, Learning, and Assessment,* 6(9), 4-37.

Bill and Melinda Gates Foundation. (2014). *Teachers knows best: What educators want from digital instructional tools*, Retrieved from https://s3.amazonaws.com/edtech-production/reports/Teachers-Know-Best_0.pdf

Clarke-Midura, J., & Dede, C. (2010). Assessment, technology, and change. *Journal Of Research on Technology In Education, 42*(3), 309-328.

Dee, T., & Jacob, B. A. (2010). Evaluating NCLB. *Education Next, 10*(3) Retrieved from https://goucher.idm.oclc.org/login?url=https://search-proquest-com.goucher.idm.oclc.org/docview/1237819643?accountid=11164

Edutopia (2012). An Introduction to Technology Integration [Video file] Retrieved from https://www.edutopia.org/video/introduction-technology-integration

Eyyam, R., & Yaratan, H. S. (2014). Impact of use of technology in mathematics lessons on student achievement and attitudes. *Social Behavior and Personality: An International Journal, 42*(1), 31-42. doi:10.2224/sbp.2014.42.0.S31

Gray, L., Thomas, N., & Lewis, L. (2010). Teachers' use of educational technology in U.S. public schools: 2009. *Education Week, 29*(31), 4. Retrieved from https://search-proquest-com.goucher.idm.oclc.org/docview/313912118

Hain, B. (2011). PARCC consortium -- A five component assessment model. *Reading*

    Today, 29(1), 24. Retrieved from

    https://goucher.idm.oclc.org/login?url=http://search.ebscohost.com.goucher.idm.oclc.org/logi

    n.aspx?direct=true&db=a9h&AN=65444365&login.asp&site=ehost-live&scope=site

Holland, J., & Holland, J. (2014). Implications of shifting technology in education. *TechTrends,*

    58(3), 16-25. doi:http://dx.doi.org.goucher.idm.oclc.org/10.1007/s11528-014-0748-3

Laborda, J. G., Sampson, D. G., Hambleton, R. K., & Guzman, E. (2015). Guest editorial:

    Technology supported assessment in formal and informal learning. *Journal of Educational*

    Technology & Society, 18(2), 1-2. Retrieved from

    https://goucher.idm.oclc.org/login?url=https://search-proquest-

    com.goucher.idm.oclc.org/docview/1683511593?accountid=11164

Maguire, K. A., Smith, D. A., Brallier, S. A., & Palm, L. J. (2010). Computer-based testing: A

    comparison of computer-based and paper-and-pencil assessment. *Academy of Educational*

    Leadership Journal, 14(4), 117-125. Retrieved from

    https://goucher.idm.oclc.org/login?url=https://search-proquest-

    com.goucher.idm.oclc.org/docview/763256555?accountid=11164

Moore, W. (2003). Facts and assumptions of assessment: Technology, the missing link. *T.H.E.*

    Journal, 30(6), 20-26. Retrieved from https://goucher.idm.oclc.org/login?url=https://search-

    proquest-com.goucher.idm.oclc.org/docview/214799721?accountid=11164

Nguyen, D. M., Hsieh, Y. J., & Allen, G. D. (2006). The impact of web-based assessment and

    practice on students' mathematics learning attitudes. *The Journal of Computers in*

    Mathematics and Science Teaching, 25(3), 251-279. Retrieved from

https://goucher.idm.oclc.org/login?url=https://search-proquest-

com.goucher.idm.oclc.org/docview/220630959?accountid=11164

Partnership for 21st Century Learning (2007). *Framework for 21st century learners.* Retrieved

from http://www.p21.org/our-work/p21-framework

Poggio, J., Glasnapp, D. R., Yang, X., & Poggio, A. J. (2005). A comparative evaluation of score

results from computerized and paper pencil mathematics testing in a large-scale state

assessment program. *The Journal of Technology, Learning and Assessment, 3*(6)

Project Tomorrow. (2014). *The new digital learning playbook: Understanding the spectrum of

students' activities and aspirations*. Retrieved from

http://tomorrow.org/speakup/pdfs/su13studentsreport.pdf

Yuan-Hsuan, L., Waxman, H., Jiun-Yu Wu, Michko, G., & Lin, G. (2013). Revisit the effect of

teaching and learning with technology. *Journal of Educational Technology Society, 16*(1),

133.