A COMPARISON FROM 2008-2015 BETWEEN MISSOURI PUBLIC SCHOOL

STUDENT COMPUTER-BASED AND PAPER-AND-PENCIL BASED HIGH

STAKES ASSESSMENTS BY RACE AND SOCIO-ECONOMIC STATUS

Dissertation Proposal

Presented to

The Faculty of the Graduate School

University of Missouri-Columbia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

By

JOSHUA A. PETERS

Dr. Timothy J. Wall, Dissertation Supervisor

July, 2016

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

A COMPARISON FROM 2008-2015 BETWEEN MISSOURI PUBLIC SCHOOL

STUDENT COMPUTER-BASED AND PAPER-AND-PENCIL BASED HIGH

STAKES ASSESSMENTS BY RACE AND SOCIO-ECONOMIC STATUS

presented by Joshua Peters, a candidate for the degree of doctor of education, and hereby certify that, in their opinion, it is worthy of acceptance.

| Professor Timothy Wall |
| Professor Carole Edmonds |
| Professor William Gordon |
| Professor Chad Brinton |

ACKNOWLEDGMENTS

TABLE OF CONTENTS

iv

List of Tables

List of Figures

A COMPARISON FROM 2008-2015 BETWEEN MISSOURI PUBLIC SCHOOL

STUDENT COMPUTER-BASED AND PAPER-AND-PENCIL BASED HIGH

STAKES ASSESSMENTS BY RACE AND SOCIO-ECONOMIC STATUS

Joshua A. Peters

Dr. Timothy Wall, Dissertation Supervisor

ABSTRACT

There is a lack of knowledge in whether there is a difference in results for students on paper and pencil high stakes assessments and computer-based high stakes assessments when considering race and/or free and reduced lunch status. The purpose of this study was to add new knowledge to this field of study by determining whether there is a difference in results for students on paper and pencil high stakes assessments and computer-based high stakes assessments when considering race and/or free and reduced lunch status. The measurements in this study included 8th grade Missouri Assessment Program mathematics exam results data, Algebra I Missouri End-of-Course Exam results data, and the math portion of the ACT results data. The research design used a static-group comparison design that utilized a quantitative, non-experimental correlational methodology. By contributing further to the existing research, the findings of this study provide educators and policy makers additional information to make decisions that inform future policy discussions regarding the use of computer based testing for high stakes assessments within the K-12 public education setting.

CHAPTER 1: INTRODUCTION TO THE STUDY

Both education and technology impact people's lives throughout the world in multiple ways. The potential exists to make lives better, particularly through education and technology, which are areas of interest to people who are passionate about how education and technology continue to evolve. Technology has transformed education in many ways and has opened up curriculum within the classroom to so many possibilities (McKnight, 2012). Chalk boards, white boards, overhead projectors, and even smart boards are being cast aside and replaced with ceiling projectors, one-to-one devices, and wireless internet for students to use. Giving every student their own technology device to use for their education provides endless potential for technology to impact learning and lives. Students are able to access information gathering from many resources without entering a library, students can learn remotely from all around the world, and students can collaborate with others to work on projects together no matter where the students are located (Lynch, 2015).

Advances in technology have also provided the opportunity to change the way the assessment process is completed. In 2012, the U.S. Department of Education had more than 10,000 fourth graders complete two 30-minute writing assignments on the computer. An analysis released to the public in December of 2015 compared the results to a paper-and-pencil version given to fourth graders in 2010. The higher performing students did better on the computer, but the average and low performing students crafted better sentences on the paper-and-pencil assessment. The analysis found that low-income, black, and Hispanic students tended to be in the lower performing group and that the higher performing students were more likely to have a computer with internet access at home (Barshay, 2016). This study sought to determine whether additional assistance

might be required within technology and the implementation of technology in educational environments to ensure equity is considered in the migration of assessment methods.

Background

Students, teachers, administrators, politicians, and almost every person in the world is impacted by education. Virtually everyone is passionate about how their life has been or can be significantly impacted by education. As the structures, policies, and procedures continue to evolve and change, nearly everyone has an opinion about what is helpful and what is detrimental toward the ultimate goal of providing the highest level of education possible to every student.

The world of education has continued to evolve as a result of technological advancements. In the late 19th century, algebra was barely introduced in schools, if at all (Roschelle, 2000). Now in the 21st century, there exists a more global economy and international business is becoming a greater opportunity for business. As global competition increases and technology amplifies the ability for communication, companies and jobs are being moved throughout the world. The education systems have begun to compete with each other all around the world rather than competing on an internal level from state to state. In a speech given by Arne Duncan, former U.S. Secretary of Education, on October 19, 2010 to the Council on Foreign Relations, he discussed the idea that things have evolved internationally from just one generation ago and that international competition in higher education as well as the job market were both growing dramatically. He stated "In practical terms, globalization means that U.S. students will have to compete throughout their careers with their peers in South Korea, Canada, China, European countries, India, and other rapidly developing states" (Duncan, 2010, para. 14). The expectation for high school students today has drastically risen with

some Missouri high school graduates being expected to take end of course exams in Algebra II (Missouri Department of Elementary and Secondary Education [DESE], 2015). Technology is constantly changing and evolving. Telephones, for example, have evolved from requiring a hard line to connect one phone to another with an operator physically plugging cables together to make those connections to people now walking around with phones that fit in their pockets. Additionally, most phones have the ability for the user to ask questions and get assistance at a push of a button regardless of their location. With the technological advances that have been made, there has been an explosion of knowledge that is now available to the public, as well as readily accessible solutions to the growing demands of the workplace (Roschelle, 2000). To meet these demands, schools are faced with analyzing resources and policies to determine how to provide the equipment and infrastructure needed to support this acceleration. The schools need to not only provide the technology, but also teach staff and students how to effectively use it (Weglinsky, 1998). Schools view technology as a tool to help increase the rigor of learning and the curriculum that students are exposed to. This would then, in theory, increase what the students would be able to produce within an assessment.

In January of 2002, President Bush signed into law the No Child Left Behind Act of 2001 (NCLB Act). This was a time of concern by the public for education in the United States. The government sought out a way to more actively engage in improving public education. "The NCLB legislation set in place requirements that reached into virtually every public school in America. It expanded the federal role in education and took particular aim at improving the educational lot of disadvantaged students" (Editorial Projects in Education Research Center, 2011, para. 2). The Act had a few key components that included required annual testing; "annual yearly progress" of student

performance on state assessments; giving of annual report cards for school districts and districts must provide the same type of information on each school within their district; and teacher certification in the core subject areas the teacher teaches in.

The Missouri Assessment Program (MAP) was already in place when NCLB Act was signed into law. The MAP was "designed to measure how well students acquire the skills and knowledge described in Missouri's Learning Standards (MLS)" (Missouri Department of Elementary and Secondary Education, 2014, p. 4), and so it already was meeting some of the requirements of NCLB Act. The NCLB Act did require additional testing to be put in place; it also required that the Missouri DESE examine the existing assessments to determine whether the assessments needed to be modified to meet the new requirements.

The state of Missouri's DESE changed its version of high stakes assessments from high school grade-level assessments to End of Course (EOC) exams in the 2008-2009 school year. The EOCs then moved fully online in the fall of 2010 (Missouri DESE, 2014). The ACT was still being given as a paper-and-pencil (P&P) assessment during this time. This identified the need for school districts in Missouri to ensure the districts had the adequate equipment in place to test the necessary students utilizing online assessments as the format.

As Missouri DESE transitioned its EOC exams from the traditional P&P administration, DESE slowly began to offer the ability to administer the EOC exams using a computer-based assessment (CBA) model. The Missouri DESE urged school districts to determine whether their current technology plan could support this transition. Districts began to examine at how the districts could fund any necessary upgrades to equipment in order to meet the technology requirements necessary to offer CBA testing.

The districts feared that all the money spent on technology could be wasted if the districts did not then contribute an equal effort in the use and integration of the technology into the instruction and curriculum (Zehr, 1997). Technology needed to replace traditional methods of instruction like taking notes and placing them in a PowerPoint; but it also needed to offer a more comprehensive approach to learning by giving options to students like incorporation of media into the PowerPoint notes. These options lead to additional budgetary concerns in terms of support and staff development that should complement these technology upgrades. Each advance made in technology comes with the potential for efficiency, enhancement, or accessibility (Roschelle, 2000). There also exists the potential for technology advancements to perpetuate any inequities that may be present (Sadker, 2002).

Theoretical Framework

The theoretical framework of this study was centered on research and theories that pertain to technology and how it affects curriculum and student learning experiences. The theory of technological determinism "claims that technology is the dominant factor in social change" (Bimber, 1990, p. 2). Beginning with the Industrial Revolution, technology has played a significant force within society. In the early history of the United States, scientists believed technology advances could help society to realize the American Dream of success and wealth. Technology was a tool to help with efficiency and effectiveness as people looked for prosperity (Smith, 1994). The foundation of the theory of technological determinism rests on Utopian determinists that believe within instructional technology "the growth and expansion of instructional technologies can lead to utopian (or nearly utopian) learning environments" (Surry, 1997, p. 4). There are also the dystopian determinists within instructional technology that seek to delay the adoption

of technology in education because the determinists "fear technology will replace teachers or fundamentally dehumanize the educational process" (Surry, 1997, p. 5). Richard Miller described technological determinism as society changing to adapt to new technology as it is developed (Bimber, 1990). This theory can easily be applied to the area of assessments within the educational environment. As the internet and computers evolved, instructional technology led to CBAs becoming available and society changing to acknowledge them and figure out how best to utilize them. Using this theory as a foundation, the way in which curriculum is developed and delivered in the classroom has continued to evolve with technology advancements. As those advancements are made within the curriculum, changes also occur with delivery methods of content and delivery methods of assessment of the curriculum.

The conceptual underpinnings supporting the research were technology in schools, CBAs, as well as test anxiety and educational testing. As technology in schools began to make advancements, CBAs have become more widely used, which prompted educators and various stakeholders to consider how this alters educational testing and the impact CBAs could impact testing anxiety. Computers are not only changing the access that educators have to curriculum or how educators can alter the student learning experiences, but computers also affect the area of assessment. The CBA provides the opportunity and means to consider how assessments in education can evolve (McDonald, 2002). Bennett (1998) discussed that for many first generation CBAs, the assessments are relatively identical to the assessments that were given prior to the integration of technology.

Today's computerized tests automate an existing process without re-conceptualizing it to realize the dramatic improvements that the

6

innovation could allow. Thus, these tests are substantively the same as those administered on paper: they measure the same skills, use the same behavioral designs, and depend primarily on the same types of tasks (Bennett, 1998, p. 3).

Bennett made the argument that the technology acts merely as a substitute rather than a transformation of the assessments. Therefore, this study viewed the area of high stakes assessment through the lens of technological determinism.

Rationale for the Study

Research shows that the results of prior studies regarding the use of technology to increase student achievement are mixed and contradictory (Attewell, 2001; Newhouse, 2011; Ricketts & Wilks, 2002). Studies have examined technology within education, but the studies largely examined the impact on instruction and less on the impact of assessment (Weglinsky, 1998). Therefore, further research is necessary in regard to technology's effect on student outcome. The relationship needs to be examined in regards to the effect on student outcomes when technology is expanded in the instructional setting within education, but also when technology is incorporated more into assessments within education (Weglinsky, 1998).

Maguire, Smith, Brallier, and Palm (2008) conducted a study that compared assessment results for the same assessment administered by P&P versus computer-based using two different control groups. The study did not examine socio-economics or race (Maguire, et. al., 2010). Escudier, Newton, Cox, Reynolds, and Odell (2011) published a study that compared assessments results for dental undergraduate students that completed both paper-and-pencil and CBAs on the same content. Once again, the study did not examine socio-economics or race and was conducted in a higher educational setting

(Escudier et al., 2011). Hargreaves, Shorrocks-Taylor, Swinnerton, Tait, and Threlfall (2010) conducted a similar study in England that had 10-year-old students take a P&P test in mathematics, and then a computer-based version of the test a month later. The authors argued that the computer-based test (CBT) had an overall positive effect on performance, but could not determine whether it was significant. Again, the study did not examine socio-economics or race and the students were taking a test that was not high-stakes (Hargreaves et. al., 2010).

Kim and Huynh (2007) published a study that compared results of computer and P&P versions of both algebra and biology EOC assessments. In this study the authors argued that more research was needed on high-stakes assessments to investigate administration mode effects within the K-12 arena. Kim and Huynh looked at gender, race, and socio-economic status for the overall sample population, but did not examine whether any subgroups had different results then the norm. In this study, students took one version of the test and then soon after took another version of the same test. Kim and Huynh determined a limitation of the study was the voluntary participation. "Schools in which technology was more accessible to students might have been more likely to participate" (Kim and Huynh, 2007, p. 25). Kim and Huynh also argued that future studies should include samples with more diverse backgrounds due to the convenience nature of the sample. The authors admitted that African Americans and Hispanics were particularly underrepresented in the sample compare to the general makeup of the entire middle/high school students in the state. Also, with the requirements of the NCLB Act to provide assessment data on sub-group performance, Kim and Huynh argued that studies need to be done to look at mode effects for these subgroups specifically.

Poggio, Glasnapp, Yang, and Poggio (2005) examined score results for 7th grade

8

students within mathematics testing in a large-scale assessment program. The study used the same assessment for both as the P&P assessment was transferred to computer. "Reasonable study controls support generalizability of these findings, but findings are limited to middle/level grade (7th grade) assessment of mathematics in a general education population" (Poggio, et. al., 2005, p. 26). This study was similar to Kim and Huynh's study in which it was voluntary and students took both versions of the assessment. Specific research looking into public school K-12 student performance on existing assessments when technology is then incorporated has been studied. However, the gap in research is still apparent when looking at high stakes assessments and considering socio-economics and race at the high school level. This study attempted to clear the present gap and help inform stakeholders to continue to make decisions about what format will be used to give high stakes assessments moving forward.

Statement of the Problem

There is a lack of knowledge in whether there is a difference in results for students on P&P high stakes assessments and computer-based high stakes assessments when considering race and/or free and reduced lunch status. Herold (2016, para. 1) asserted that "In 2015-16, for the first time, more state standardized tests for the elementary and middle grades will be administered via technology than by paper and pencil." Herold went on to state that both academics and parents alike have concerns that achievement gaps may be widened as student have unequal access to and use of technology. It is a reasonable belief to think that the increase of technologies use within education will help to level the playing field for current racial, gender, and geographic divides (Noeth, 2004). This cannot be made as an absolute, however, without data to support the belief. Is there a difference in results for students on P&P high stakes

9

assessments and computer-based high stakes assessments when considering free and reduced lunch status and/or race? This study looked to close the knowledge gap in the research and examine whether there is an increased effect on the difference between the 8th grade MAP assessment and the Algebra I EOC test results in comparison to the difference between the 8th grade MAP assessment and ACT test results when taking into consideration race and socio-economics of the student.

Purpose of the Study

The purpose of this study was to add new knowledge to this field of study by determining whether there is a difference in results for students on P&P high stakes assessments and computer-based high stakes assessments when considering free and reduced lunch status and/or race. To accomplish this task, this study took into consideration the established relationships of a P&P assessment and the use of an online assessment as seen in Figure 1 and Table 1. By contributing further to the existing research, the findings of this study may provide educators and policy makers additional information to make decisions that inform future policy discussions regarding the use of CBT for high stakes assessments within the K-12 public education setting.
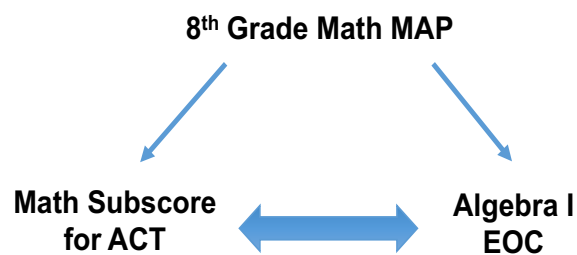
Figure 1. Math High Stakes Assessments

Table 1. *Assessment Format*

	Paper-and-Pencil	CBA
8th Grade Math MAP	X	
Algebra I EOC		X
Math Subscore for ACT	X	

Research Questions

The research questions and null hypotheses that guided this study were:

Research Question 1

What are the descriptive statistics for student performance on the 8th grade math MAP test, the Algebra I EOC examinations, and the math portion of the ACT broken out by the following subgroups?

By Race:

a. Minority and Not Minority

By Free/Reduced Lunch Status:

a. Qualifiers and Non-qualifiers

$H_0 1$. There is no difference in descriptive statistics of student performance on the 8th grade math MAP test, the Algebra I EOC examinations, and the math portion of the ACT when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

Research Question 2

Is there a difference in the relationship between student performance on the 8th grade math MAP examinations when broken out by the following subgroups?

By Race:

a. Minority and Not Minority

By Free/Reduced Lunch Status:

a. Qualifiers and Non-qualifiers

H_02. There is no difference in the relationship between student performance on the 8th grade math MAP examinations when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

Research Question 3

Is there a difference in the relationship between student performance on the Algebra I EOC examinations when broken out by the following subgroups?

By Race:

a. Minority and Not Minority

By Free/Reduced Lunch Status:

b. Qualifiers and Non-qualifiers

H_03. There is no difference in the relationship between student performance on the Algebra I EOC examinations when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

Research Question 4

Is there a difference in the relationship between student performance on the math portion of the ACT when broken out by the following subgroups?

By Race:

a. Minority and Not Minority

By Free/Reduced Lunch Status:

a. Qualifiers and Non-qualifiers

H_04. There is no difference in the relationship between student performance on the Algebra I EOC examinations when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

Research Question 5

What is the difference in the measurement of strength of association between the 8th grade math MAP examination and the Algebra I EOC examination compared to the 8th grade math MAP examination and the math portion of the ACT when broken out by the following subgroups?

By Race:

a. Minority and Not Minority

By Free/Reduced Lunch Status:

b. Qualifiers and Non-qualifiers

$H_0 5$. There is no difference in the measurement of strength of association between the 8th grade math MAP examination and the Algebra I EOC examination compared to the 8th grade math MAP examination and the math portion of the ACT when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

Limitations of the Study

This study was limited in a number of ways. First, this study included only students that went on to take the ACT test. Students that do not plan on attending a 4-year college will take different assessments. Additionally, not all students that plan to attend a 4-year college will take the ACT, but some will take the SAT instead. This study was limited to high schools in the Kansas City, Missouri area. The largest percentage of students going onto a 4-year college in this area takes the ACT assessment.

Additional limitations were that the EOC exams are used only in Missouri and different types of assessments in different states may have different results. The tests are

created independently and different steps may be taken to transition assessments to an online format. Another study limitation was that students may have been subjected to varying test preparation models that may impact the results outside of the variables being studied. Also, students were sampled from schools that had a 1:1 technology environment within the school. Schools that did not have a 1:1 technology environment were not used within the study. The final limitation of the study was that results of this study described only the population that took these assessments during the 3-year period (between 2013 and 2015) of high school graduates examined.

Definition of Key Terms

The following terms are presented to clarify their use within the context of the study.

ACT: The ACT Assessment is a curriculum- and standards-based educational and career planning tool that assesses students' academic readiness for college (ACT, 2014).

Computer-Based Assessment (CBA): The use of digital tools for assessment-related activity (ProProfs Quizmaker, 2014).

Curriculum: The lessons and academic content taught in a school or in a specific course or program (Great Schools Partnership, 2014).

End of Course (EOC): An EOC is a yearly standards-based grade-level assessment that measures specific skills defined by the state of Missouri (Missouri DESE, 2014).

High-Stakes Tests: Any test used to make important decisions about students, educators, schools, or districts, most commonly for the purpose of accountability—i.e., the attempt by federal, state, or local government agencies and school administrators to ensure that students are enrolled in effective schools and being taught by effective

teachers (Great Schools Partnership, 2014).

One to One: Programs that provide all students in a school, district, or state with their own laptop, netbook, tablet computer, or other mobile-computing device. *One-to-one* refers to one computer for every student (Great Schools Partnership, 2014).

Paper-and-Pencil (P&P) Based Assessment: Is a general group of assessment tools in which candidates read questions and respond in writing (Public Service Commission of Canada, 2011).

School Technology: For the purposes of this study, school technology is defined as computer equipment that includes computers, software, and other equipment and infrastructure that is used in a classroom (Technology in Schools Task Force, 2002).

Student Learning Experiences: Any interaction, course, program, or other experience in which learning takes place, whether it occurs in traditional academic settings (schools, classrooms) or nontraditional settings (outside-of-school locations, outdoor environments), or whether it includes traditional educational interactions (students learning from teachers and professors) or nontraditional interactions (students learning through games and interactive software applications) (Great Schools Partnership, 2014).

Summary

This chapter discussed a non-experimental, quantitative research study, and introduced a brief background of the impact of technology on education as well as the need to evaluate the relationship between computer-based and P&P-based high stakes assessments when considering race, and socio-economic status. Additionally, this chapter defined the conceptual underpinning of the study, the purpose of the study, and the research questions. The study hypotheses were provided. Finally, this chapter listed the

limitations of the study and the defined key terms. The next chapter elaborates on the

conceptual underpinnings and review relevant literature that will inform the study.

CHAPTER 2: REVIEW OF RELATED LITERATURE

Background

Students, teachers, administrators, politicians, and almost every person in the world is impacted by education. Virtually everyone is passionate about how their life has been or can be significantly impacted by education. As the structures, policies and procedures continue to evolve and change, nearly everyone has an opinion about what is helpful and what is detrimental toward the ultimate goal of providing the highest level of education possible to every student.

The world of education has continued to evolve as a result of technological advancements. In the late 19th century, algebra was barely introduced in schools if at all (Roschelle, 2000). Now in the 21st century there exists a more global economy and international business is becoming a greater opportunity for business. As global competition increases and technology amplifies the ability for communication, companies and jobs are being moved throughout the world. The education systems have begun to compete with each other all around the world rather than competing on an internal level from state to state. In a speech given by Arne Duncan, former U.S. Secretary of Education, on October 19, 2010 to the Council on Foreign Relations, he discussed the idea that things have evolved internationally from just one generation ago and that international competition in higher education as well as the job market were both growing dramatically. He stated "In practical terms, globalization means that U.S. students will have to compete throughout their careers with their peers in South Korea, Canada, China, European countries, India, and other rapidly developing states" (Duncan, 2010, para. 14). The expectation for high school students today has drastically risen with some Missouri high school graduates being expected to take end of course exams in

Algebra II (Missouri Department of Elementary and Secondary Education [DESE], 2015). Technology is constantly changing and evolving. Telephones, for example, have evolved from requiring a hard line to connect one phone to another with an operator physically plugging cables together to make those connections to people now walking around with phones that fit in their pockets. Additionally, most phones have the ability for the user to ask questions and get assistance at a push of a button regardless of their location. With the technological advances that have been made, there has been an explosion of knowledge that is now available to the public, as well as readily accessible solutions to the growing demands of the workplace (Roschelle, 2000). To meet these demands, schools are faced with analyzing resources and policies to determine how to provide the equipment and infrastructure needed to support this acceleration. The schools need to not only provide the technology, but also to teach staff and students how to effectively use it (Weglinsky, 1998). Schools view technology as a tool to help increase the rigor of learning and the curriculum that students are exposed to. This would then, in theory, increase what the students would be able to produce within an assessment.

In January of 2002, President Bush signed into law the No Child Left Behind Act of 2001 (NCLB). This was a time of concern by the public for education in the United States. The government sought out a way to more actively engage in improving public education. "The NCLB legislation set in place requirements that reached into virtually every public school in America. It expanded the federal role in education and took particular aim at improving the educational lot of disadvantaged students" (Editorial Projects in Education Research Center, 2011, para. 2). The Act had a few key components that included required annual testing; "annual yearly progress" of student performance on state assessments; giving of annual report cards for school districts and

then districts must provide the same type of information on each school within their district; and teacher certification in core subject areas in which the teachers were teaching.

The Missouri Assessment Program (MAP) was already in place when NCLB Act was signed into law. The MAP was "designed to measure how well students acquire the skills and knowledge described in Missouri's Learning Standards (MLS)" (Missouri Department of Elementary and Secondary Education, 2014, p. 4) and it was already meeting some of the requirements of NCLB Act. The Act required additional testing to be put in place; it also required that the Missouri DESE examine the existing assessments to determine whether the assessments needed to be modified to meet the new requirements.

The state of Missouri's DESE changed its version of high stakes assessments from high school grade-level assessments to End of Course (EOC) exams in the 2008-2009 school year. The EOCs then moved fully online in the fall of 2010 (Missouri DESE, 2014). The ACT was still being given as a paper-and-pencil (P&P) assessment during this time. This identified the need for school districts in Missouri to ensure the districts had the adequate equipment in place to test the necessary students utilizing online assessments as the format.

As Missouri DESE transitioned its EOC exams from the traditional P&P administration, DESE slowly began to offer the ability to administer the EOC exams using a computer-based assessment (CBA) model. The Missouri DESE urged school districts to determine whether their current technology plan could support this transition. Districts began to examine at how the districts could fund any necessary upgrades to equipment in order to meet the technology requirements necessary to offer CBA testing.

The districts feared that all the money spent on technology could be wasted if the districts did not then contribute an equal effort in the use and integration of the technology into the instruction and curriculum (Zehr, 1997). Technology needed to replace traditional methods of instruction like taking notes and placing them in a PowerPoint; but it also needed to offer a more comprehensive approach to learning by giving options to students like incorporation of media into the PowerPoint notes. These options lead to additional budgetary concerns in terms of support and staff development that should complement these technology upgrades. Each advance made in technology comes with the potential for efficiency, enhancement, or accessibility (Roschelle, 2000). There also exists the potential for technology advancements to perpetuate any inequities that may be present (Sadker, 2002).

This chapter examines the research relating to technology in schools in both past and present. This review also presents research on high stakes assessments, and concludes with an examination of research pertaining to assessment challenges including anxiety.

Technology in Schools

Technology has existed in schools for decades. The implementation of technology is not easily accomplished and can present many challenges. Even with procedural challenges and strain placed on funding and resources, leaders in the education sector still recognize the significance of technology as it continues to evolve. Technology is typically included when educators discuss ways to help prepare students to be successful in the future (Celano, 2010).

History of Technology Implementation

Implementation of computer use in education had a slow start, mainly because

computers were a costly investment and the return on that investment had not been explored enough to warrant the significant budget shifts necessary for school districts; specifically, schools would need to finance the hardware and software, as well as wiring and the necessary infrastructure to support the technology. Few schools were online in the 1980s, but according to the National Center for Education Statistics (NCES), by the year 2000, almost all schools were online (Sadker, 2002). Congress recognized the high cost to bring computer technology into the schools and so Congress passed the Telecommunications Reform Act (1996), which established a discounted cost for schools. This was also known as the E-rate (education rate) and was used to increase access to technology in schools. The problem was that the hardware and the software were also considerable expenses for schools (Sadker, 2002).

As schools found ways to afford computers, the issue became the lack of a computer for each student, which hindered teachers in using the computers as resources in their classrooms. Computers slowly transitioned from being housed in computer labs to being a regular fixture in everyday classrooms. Today we see more and more schools providing individual computer devices for students (Heinich, 2002). This initiative has been called 1:1 and provides either a means for students to access a device while at school or assignment of a computer to each student—similar to being assigned a textbook for the school year—that is for use both at school and at home.

As the numbers of computers have grown in schools, the number of computers also have grown in businesses and homes. The increase in exposure to computers for teachers and students also brings advancements with the technology itself, some of which make technology more accessible, easier to use, and less intimidating and more affordable (Heinich, 2002). Teachers and students became more confident and

comfortable when using technology as both groups have more exposure to technology both inside and outside of school.

Access to Technology

While teacher and student exposure to technology had increased, there has not been a correlation to the increase in the availability of technology. Wealthier communities and wealthier schools have seen a quicker and greater access to the internet and thus a higher rate of access to various educational resources available to them from outside their immediate community (Sadker, 2002). As much as socio-economics impacted the pace at which students and teachers gained exposure, geographical location played a significant impact as well. Rural schools had less access to the costly and much faster fiber optic cables resulting in the internet being slow and less reliable. This made implementation of online classes more difficult, as well as impacted what teachers could do with curricular materials to support their instruction that could be found online (Sadker, 2002). Districts that were more rural and had less tax-based income were also the same communities that incurred higher costs to integrate the same level of technology and access as other more urban communities.

The amount and type of access to technology that students and teachers had a significant impact on how easily and effectively technology was incorporated into the educational environment. Students were negatively impacted when the students had limited access to technology in school and at home, while other students with access pulled ahead academically (Garcia, 2001). To have an equitable impact between schools, a strong technological infrastructure was necessary both around the school as well as in the school. It takes teachers considerable time to plan for technology use in the classroom; the teachers do not want to use time to plan for something that is not reliable

and may not be able to be used in the end (Weglinsky, 1998). In schools where technology may be present, but a reliable and substantial infrastructure doesn't exist, implementation of the technology will undoubtedly suffer. Teachers do not want to plan and count on a resource as part of their lesson only to have issues that render the lesson incomplete.

As technology in schools became more and more prevalent, research was conducted to examine the potential impact on student performance. James Kulik conducted a meta-analysis of over 500 individual research studies and determined that overall, the students that used computer-based instruction scored higher than students that did not (Schacter, 1999, p. 4). In 2011, the National Assessment of Educational Progress transitioned their P&P writing assessment over to a CBA model. Striving to align their assessment format to reflect the same format that students were expected to write to, it was determined that the CBA was more aligned with the role technology played as students move through their education and subsequently into the workforce (NAEP, 2010). The need for training for teachers on the technology and the possibilities for implementation within the curriculum and classroom only grew. It became apparent that teachers would need training on how to use computers and technology to avoid a negative impact on their instruction and curriculum (Earle, 2002).

Educational leaders continued to examine how educators could best support their students and staff through relevant professional development and a significant allotment of time and resources to learn. Technologically savvy staff could have a significant impact on student preparation for potential career and/or college choices after public education. Educators were challenged with providing opportunities for students to use technology that provided them global access in a way in which the educators would use it

in the workforce (Kist, 2013). This required a continual evaluation of the efficiency and effectiveness of procedures and resources.

Equity of Technology

Through many programs and resources, schools have been a place for students to be given technological equity with other students. Schools typically have been one platform used to even out disparities among students and technology, which presents a challenge to schools to accomplish (Heinich, 2002). To remain consistent with other components of school, many districts began to look at how to provide equity within technology as well.

School districts considered potential ways to get the same technology device to every student regardless of their family's financial situation. Depending on finances at home, students were coming to school with different levels of technology skills that placed the students with more access at home with an advantage over the students that did not have the access or practice working in a technology environment. Students that lagged behind had the potential to fall even further behind as the students focused on learning the technology and not on the content, while the others students could focus right away on content (Celano, 2010). School districts slowly started making decisions to use a 1:1 platform in terms of computers, where every student would have access to their own technological device to use regularly within their education. Through consistent access to technology, students began to have regular exposure and practice with computers (Dean, 2012). The more a student has regular access to a tool or setting, the more comfortable and confident the students become with utilizing it. By providing students with regular opportunities to develop experience with technology, the computer becomes less an obstacle to learning, and more an additional tool to enable additional

content to be learned.

When every student in the classroom had the same access to technology within a classroom, the impact on instruction expanded as teachers no longer had to plan for the students that may or may not have access to technology as the schoolwork is completed. Students continued to be able to access curriculum through books, but the regular use of computers in the classroom helped them to be acclimate to this type of classroom setting (Celano, 2010). The transformation of textbooks to a digital format provides multiple ways of teaching a specific content and gives the teacher the ability to more individualize the instruction to the student's needs. The digital textbook also helps students to gain access when the students are not in school or miss a class for any reason (Horton & Lovitt, 1994). The student need for their teachers increased as well; students need their teachers to know the content as well as have a foundational knowledge of technology and how to use it in order to teach the student how to use it.

> As the National Council of Teachers of English suggests, literacy is much
> more than simply reading and writing texts. The organization's position
> statement (n.d) now defines 21st century literacies as including
> 'proficiency with the tools of technology,' an ability to 'manage, analyze,
> and synthesize multiple streams of simultaneous information,' an ability to
> 'design and share information for global communities to meet a variety of
> purposes,' and more. (Richardson, 2013, p. 13)

By providing a device for every student, schools are able to continue to make advances in preparing students for their futures after high school graduation.

High Stakes Assessments

Assessments within education have emerged as a critical component to help guide

reform and change. High stakes assessments are being used to gauge a school's effectiveness and success. (Poggio, 2005). Tests become high stakes when the results have a potential impact on an individual educationally or occupationally. The tests become high stakes for schools when used to determine performance levels (Stobart, 2012). These high stakes assessments tend to be used by state organizations and politicians to compare schools across a state. The provisions of NCLB required reporting of statewide test results for all major subgroups (e.g., students with disabilities, English Language Learners, students that qualified for free and reduced lunch, and students that were in various ethnicity groups). Schools also were required to report overall student performance (Kim and Huynh, 2010).

Computer-based Assessments

With the addition of technology and computers to the school environment comes the ability to transform curriculum and instruction, and the opportunity to transform the world of assessments. The use of online assessments will continue to expand as the integration of technology in schools expands across the United States (Ryan, 2014). Computerized assessments can promote and support the use of technology within the classroom, but the assessments can also provide some advantages, some of which include faster turn-around of grading and results, as well as more flexibility for test items including media imbedded questioning (Willhoft, 2012). Additionally, the administration of testing to students that requires accommodations due to disabilities could gain greater flexibility and schools save money through reduced printing costs. There is also the ability to increase testing security measures when using computerized assessments (Poggio et al., 2005). School personnel would potentially have data sooner to utilize to make decisions of how to better support students that may need remediation and to meet

the assessment reporting requirements of the NCLB Act.

This format also requires additional planning outside of the test design process. Educators would be doing a disservice to their students if the educators did not recognize the time and effort that this transformation requires. Instruction and assessment formats need to be aligned so that students utilize the same format to learn the curriculum as the format that the students will be required to use to demonstrate their mastery of the content and application of the materials. Students do not want to learn using P&P and then be required to take an assessment on a computer (Ryan, 2014). Educators also need to recognize the potential implications and impact on student performance that this change has the potential to cause.

A student that is not familiar with technology may be inhibited by the use of a computer while trying to demonstrate the same level of mastery as the the student would be able to do in a P&P environment that the student is used to. If a test needs to be altered or a student needs to scroll down a screen for a question, then administration mode effects increase according to the findings of Bridgeman, Lennin & Jackenthal (2003). The different formats might produce different scores. If there were different scores, then the validity of the various test formats would need to demonstrate the same rank order regardless of the format (McDonald, 2002).

The Missouri DESE began to move the high school portion of the MAP to an online format. The name of the high school version of these assessments was changed to End-of-Course Exams (EOC). During the 2008-2009 school year, some Missouri school districts began administering the EOC assessments in Algebra I online on a voluntary basis. Students would take the Algebra I EOC when the student completed the course. The following year, all EOC exams were given online at the high school level (Missouri

DESE, 2015). The district selected in this study did choose to have their students take the assessment online during the 2008-2009 school year.

Paper and Pencil Assessments

Assessments have been given in education long before the addition of the internet and online assessments ever became an option. Before the addition of CBAS, high stakes assessments were given using paper and pencil. The push to start moving away from the traditional format really started to make progress when the United States Congress passed the NCLB Act in 2001 (ESEA, 2001), which set accountability measures for schools that was directly connected with federal funding (Lee &Reeves, 2012).

The elementary portion of the MAP were not called EOCs, but rather the assessments have been titled the MAP Grade-Level Assessments. These assessments are given for mathematics for grades 3-8. It was not until the 2014-2015 school year that the MAP Grade-Level Assessments were administered fully online for the first time (MoDESE, 2014).

The ACT is a high-stakes standardized assessment that is used by many colleges as an entrance exam. It is currently accepted by all four-year colleges and universities in the United States. It is administered six times during the year in the United States. The ACT has officially begun administering computer-based exams to select schools that participate in State and District testing. The online option will become increasingly more available starting in the Spring of 2016 (ACT, 2014). On January 14, 2014, The Missouri State Board of Education approved a measure to administer the ACT to all 11th grade students beginning in the 2014-15 school year (MoDESE, 2015). Prior to this, the option to take the ACT was voluntary for students including those within the school district studied.

Assessment Challenges

Testing Bias

Many high stakes assessments are made up of standardized tests. These tests have been studied to determine whether any biases occur within the test design. "On K-12 achievement tests and college entrance exams, lower income students, as well as black and Latino students, consistently score below privileged white and Asian students" (Rooks, 2012, para. 3). As these assessments are used in ways that have potential significant impact on students and organizations, then the understanding of how the tests may impact students by race and socio-economic status is important (Whiting & Ford, 2009). Jaschik talks about a study done in 2010 that replicated the findings of a similar study done in 2003. Both studies found that there were differences in results based on socio economics as well as race when looking at the Scholastic Aptitude Test (SAT) exam scores (Jaschik, 2010). Test bias is one portion to consider, but then one also must consider test fairness as well. When the results of these assessments start to impact opportunities for these students due to test bias, then it is important to consider if they change from paper and pencil to computer-based has an even bigger impact on these identifiable groups due to test fairness.

Testing Anxiety

Schools administer all types of assessments throughout the school year. At various time throughout the school year, benchmarks are used to gauge where the students are in their understanding and to assess the progress being made. These benchmark tests are often given in mathematics, science, English language arts, and social studies. The tests are aligned with college and career readiness expectations. Educators use these results to assess whether students have the skills and knowledge

needed to be successful in the real-world (Smith, 2014). Teachers can then use the results from these assessments for further planning and to know where their students are at in regards to mastery of standards.

The use of assessments within the classroom is a regular occurrence, but the exposure to high-stakes testing is something a student experiences less often. Standardized tests are commonly the format for high-stakes testing in education. Standardized tests have procedures and settings that are set to ensure that all students no matter the time or place are getting the test administered to them in the same way (Carey, 1994). These types of tests can be used across schools as well as districts. Achievement tests are the most common standardized test and are designed to see how much a student has learned (Woolfolk, 2001). The tests can also be given across wide groups within the educational world. These tests become high-stakes testing when the results begin to be used by school officials to make large decisions based on results in regards to things like technology and budgets (Woolfolk, 2001). Wang, Jiao, Young, Brooks, and Olson (2008) pointed out that the stakes for testing has increased due to the requirements set forth in the implementation of the NBLC Act.

If the technology does not work effectively during assessments, then anxiety is heightened and the format that the students have been prepared for needs to be changed. It has been found that the more anxious a person is directly influences their performance on a computer (Brosnan, 1998). When the students are tasked with taking an assessment in a format that the student does not use often, their anxiety is likely to be even higher due to their lack of familiarity. By embedding online assessments as part of the regular school day, this format no longer feels like a special event and thus lowers student anxiety (Ryan, 2014).

Testing Comfort

Students must be given the opportunity to become comfortable with technology through exposure to it on a consistent and frequent basis to ensure that the technology does not inhibit or stall the student's performance. The more practice a student has, the more comfortable that student is with the format and environment, thus impacting the student's overall experience (Woolfolk, 2001). Further, this would help to ensure the assessment is truly measuring the student's knowledge in the content area being assessed. As high stakes assessments continue to shift to online assessment formats, more and more school districts are trying to provide a computer for every student to use on a regular basis. The comfort level at the time of assessment is then much higher as the students already have confidence and familiarity with the device the student is taking the test on. The students understand what the assessment looks like and feels like and how better to adequately prepare themselves for that experience. Additionally, students score better when training has been provided on study skills as well as test taking strategies (Garcia, 2001). By providing regular use of technology within the classroom for both instruction and assessment, the impact of the assessment format becomes less significant.

The more exposure a student has to something, the more comfortable that student becomes. Students that have been exposed to the content, structure, or format more than other students, then, potentially have an advantage going into the test (Woolfolk, 2001). This is especially true when the tests are administered in a format that some students are less familiar with because the students have not had the same level of exposure. A frequent assumption is that all students have an equal chance at scoring well on a test because all have all been exposed to the same material (Garcia, 2001). Different students learn at different speed and in different ways. Unfortunately, there is also an assumption

that every student has had the same level of exposure to the test structures, which is not always the case either.

There are various strategies and training that can help make a difference in student scores. High-stakes testing does not always necessarily conduct the same type of assessment. The SAT and ACT tests are used by colleges and universities all over the country. These two tests are considered norm-referenced tests because the tests compare student results to other student results from all around the country (Sadker, 2002). These tests help to predict how well a student might do in college (Woolfolk, 2001). These schools used these results to make big decisions about student's futures as the schools accept or deny applications to attend their school. Objective-referenced tests are tests that are more specifically targeted and measure whether a student has mastered a designated body of knowledge (Sadker, 2002). This is the type of test that Missouri uses as their EOC exam.

Summary

The need for research in regards to the affect that computer-based assessments has on student scores is clear. High-stakes assessments have become more widely used and can have a large impact on a student and school district. Research showed that the results of prior studies regarding the use of technology to increase student achievement are mixed and contradictory (Attewell, 2001; Newhouse, 2011; Ricketts & Wilks, 2002). Studies have examined technology within education, but the studies largely examined the impact on instruction and less on the impact of assessment (Weglinsky, 1998). Therefore, further research is necessary in regard to technology's effect on student outcome.

This chapter expanded on three key components supporting the research. First, a historical journey of the progression of technology within schools was explored. Second,

an explanation of high-stakes assessments including computer-based assessments and paper-and-pencil assessments was offered. Third, the various challenges within assessments were presented including text anxiety and the myriad of formats and types of educational testing. The next chapter explores the research design and methodology of the study.

CHAPTER 3: METHODOLOGY

Introduction

This chapter on research methodology is divided into subsections. First, the statement of the problem is offered. Second, the purpose of the study is described. Next, the research questions and hypothesis are provided. Fourth, the population and sample are presented, Fifth, the research design, variables, grouping factors, data collection, privacy and security measures, and data analysis are discussed. Finally, a chapter summary is provided.

Statement of the Problem

There is a lack of knowledge in the research, which would be beneficial to examine the difference between high stakes assessments when one test has been transitioned to a computer-based assessment (CBA) model within K-12 public education settings. Herold asserted that "In 2015-16, for the first time, more state standardized tests for the elementary and middle grades will be administered via technology than by paper and pencil" (Herold, 2016, para. 1). Herold (2016) went on to state that both academics and parents alike have concerns that achievement gaps may be widened as student have unequal access to and use of technology. It is a reasonable belief to think that the increase of technologies use within education might help to level the playing field for current racial, gender, and geographic divides (Noeth, 2004). This cannot be made as an absolute, however, without data to support the belief. Is there a difference in results for students on paper-and-pencil (P&P) high stakes assessments and computer-based high stakes assessments when considering free and reduced lunch status and/or race? This study looked to close the knowledge gap in the research and examined whether there was an increased effect on the difference between the 8th grade Missouri Assessment Program

(MAP) assessment and the Algebra I End of Course (EOC) test results in comparison to the difference between the 8th grade MAP assessment and ACT test results when taking into consideration race and socio-economics of the student.

Purpose of the Study

The purpose of this study was to add new knowledge to this field of study by determining whether there is a difference in results for students on P&P high stakes assessments and computer-based high stakes assessments when considering free and reduced lunch status and/or race. To accomplish this task, this study took into consideration the established relationships of a P&P assessment and the use of an online assessment using the same students. By contributing further to the existing research, the findings of this study provided educators and policy makers additional information to make decisions that inform future policy discussions regarding the use of CBT for high stakes assessments within the K-12 public education setting.
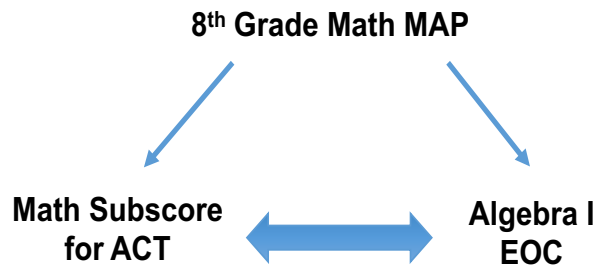
Figure 2. Math High Stakes Assessments

Table 2. *Assessment Format*

	Paper-and-Pencil	CBA
8th Grade Math MAP	X	
Algebra I EOC		X
Math Subscore for ACT	X	

Research Questions

The research questions and null hypotheses guiding this study were:

Research Question 1

What are the descriptive statistics for student performance on the 8th grade math MAP test, the Algebra I EOC examinations, and the math portion of the ACT broken out by the following subgroups?

By Race:

a. Minority and Not Minority

By Free/Reduced Lunch Status:

a. Qualifiers and Non-qualifiers

$H_0 1$. There is no difference in descriptive statistics of student performance on the 8th grade math MAP test, the Algebra I EOC examinations, and the math portion of the ACT when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

Research Question 2

Is there a difference in the relationship between student performance on the 8th grade math MAP examinations when broken out by the following subgroups?

By Race:

a. Minority and Not Minority

By Free/Reduced Lunch Status:

a. Qualifiers and Non-qualifiers

$H_0 2$. There is no difference in the relationship between student performance on the 8th grade math MAP examinations when considering race (minority and not minority)

and free/reduced lunch status (non-qualifier and qualifier).

Research Question 3

Is there a difference in the relationship between student performance on the Algebra I EOC examinations when broken out by the following subgroups?

By Race:

a. Minority and Not Minority

By Free/Reduced Lunch Status:

b. Qualifiers and Non-qualifiers

$H_0 3$. There is no difference in the relationship between student performance on the Algebra I EOC examinations when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

Research Question 4

Is there a difference in the relationship between student performance on the math portion of the ACT when broken out by the following subgroups?

By Race:

b. Minority and Not Minority

By Free/Reduced Lunch Status:

a. Qualifiers and Non-qualifiers

$H_0 4$. There is no difference in the relationship between student performance on the Algebra I EOC examinations when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

Research Question 5

What is the difference in the measurement of strength of association between the 8th grade math MAP examination and the Algebra I EOC examination compared to the

8th grade math MAP examination and the math portion of the ACT when broken out by

the following subgroups?

By Race:

a. Minority and Not Minority

By Free/Reduced Lunch Status:

b. Qualifiers and Non-qualifiers

$H_0$5. There is no difference in the measurement of strength of association between

the 8th grade math MAP examination and the Algebra I EOC examination compared to

the 8th grade math MAP examination and the math portion of the ACT when considering

race (minority and not minority) and free/reduced lunch status (non-qualifier and

qualifier).

Population and Sample

The selected research site, School District A, was located in the state of Missouri,

in the suburbs of a major metropolitan area. C High School had its first students in the

2008-2009 school year with their first graduating class being in 2010.

Table 3. *Demographic Information*

School District A	2008	2009	2010	2011	2012	2013	2014
Total Enrollment	17,552	17,955	18,288	18,530	18,674	18,928	19,199
White Percent	71.7	70.5	69.7	65.9	64.8	64.3	63.6
Free/Reduced Lunch (FTE) Percent	38.7	40.9	43.8	46	47.6	48.7	49.6
School: A HIGH							
Total Enrollment	1,812	1,614	1,562	1,507	1,482	1,440	1,505
White Percent	63.5	60.5	56	53.3	50.9	49.2	49.6
Free/Reduced Lunch (FTE) Percent	40.8	43.4	46.5	50.6	52.9	55.2	55.8
School: B HIGH							
Total Enrollment	2122	1530	1314	1351	1382	1383	1413
White Percent	81.3	77.6	74.7	73.3	70	67.8	66.8
Free/Reduced Lunch (FTE) Percent	17.1	27.9	35.3	39.1	39.5	41.6	43.1
School: C HIGH							

Total Enrollment		947	1295	1347	1377	1377	1443
White Percent		80.6	80.4	79.8	78.6	77.7	76.2
Free/Reduced Lunch (FTE) Percent		15.9	16.9	17.4	18.7	17.7	19.1
School: D HIGH							
Total Enrollment	1636	1593	1516	1430	1346	1310	1325
White Percent	74.3	74	72.6	67.4	64.2	64.5	61.4
Free/Reduced Lunch (FTE) Percent	36.8	37.6	41.5	46.2	48	49.2	55.5

(DESE, mcds.dese.mo.gov, 2014)

In 2010 for example, there were 1389 graduates in the district with 57.88% taking

the ACT (see Appendix B, ACT Results). This should then provide approximately 800

students for the sample size for that school year prior to pulling out students that did not

have an 8th grade math MAP score and an Algebra I EOC score. The students were

selected from the four high schools from School District A because the district is one of

the largest school districts in the state of Missouri, the district's student population

represented a wide sampling of diversity and socio-economic status.

This school district was picked because the four high schools within the district

spanned from urban to suburban locations and had a wide variance of free and reduced

and white versus non-white. The study was limited to the four high schools within this

district because all four high schools were exposed to 1:1 at the same time and with the

same devices, the schools had the same curriculum guides, and testing plans. All of the

high school students within School District A were given a laptop computer that the

students would use and take home with them during the school year for all four years of

high school.

Research Design

The MAP Overview

The MAP measures how well students acquire essential knowledge and skills

described in Missouri's Learning Standards (Missouri DESE, 2014a). The MAP also provides information to districts and the state regarding the overall quality of education being provided. Standardized assessments under this program are given in grades 3-12 via a variety of K-8 grade level tests and 9-12 subject area tests. For all assessments, student performance is reported in terms of four levels of performance (or achievement) with *Advanced* and *Proficient* being the highest two levels and *Basic* and *Below Basic* being the lowest two. These performance levels are used to calculate approximately 67% of a school district's accreditation score on standardized assessment data (Missouri DESE, 2014a).

Algebra I EOC Exam

For this study the researcher used achievement data from the EOC assessment taken by the student. This assessment is generally taken upon completion of 9th grade level Algebra courses and is designed to assess a student's progress toward the Missouri Learning Standards in which the student received instruction on during the Algebra course (Missouri DESE, 2014).

Grade 8 Math MAP Assessment

For this study, the researcher used achievement data from the grade 8 mathematics MAP assessment. This assessment is taken at the end of a student's 8th grade year.

ACT Assessment

For this study, the researcher used achievement data from the ACT. This assessment is typically taken during a student's 11th or 12th grade year.

The research design used a static-group comparison design that utilizes a quantitative, non-experimental, correlational methodology. Quantitative studies allow

researchers to statistically analyze numerical data for reliability, significance, and correlation by using descriptive statistics (Creswell, 2009). The study used a static-group comparison design (Fraenkel & Wallen, 2003). A non-experimental correlational study was chosen because the independent variables were not manipulated at any point in the study (Creswell, 2009). This study did not attempt to establish a causal relationship; therefore, correlational methods were used to establish the degree to which the independent and dependent variables are related (Creswell, 2009).

The study used an analysis of covariance (ANCOVA) to evaluate difference among group means. Including covariates in analysis of variance (ANOVA) helps remove any bias of variables that may be unmeasured and it reduces the error variance (Field, 2009). This allowed for an analysis of data utilizing both subsections of data gathered (i.e. Caucasian, Race, etc.). McDonald suggested using ANCOVA to compare two or more regression lines to each other. This helped to determine if the lines are different in either slope or intercept (McDonald, 2014).

The researcher established the p-value at .05 and rejected the null hypothesis if the p-value was smaller than or equal to .05. If the p-value was rejected, then the researcher pursued post-hoc tests and the results were displayed and analyzed. The analysis utilized Pearson's-r product moment coefficient of correlation, ANCOVA, measures of centrality, standard deviations, graphs, figures, and scatter plots to present the findings and facilitate the analysis of the results (Field, 2009).

Variables

Independent Variables in the Study

Three independent variables or factors were examined in this study: (a) Race: minority, and not minority, (b) free/reduced lunch status: Qualified and non-qualified,

and (c) Assessment format: P&P Based and CBA.

Dependent Variables in the Study

There were three dependent variables in the study: (a) the Algebra I EOC score,
(b) the 8th grade math MAP score, and (c) the math subsection of the ACT score. The
following table lists the dependent and independent variables in the study.

Table 4. *Independent and Dependent Variables Explored in the Study by Data Type*

Independent Variables (Factors)		Data Type
a. Race		
	a. Students From Minority Background	Binary
	b. Students Not From Minority Background	Binary
b. Free/Reduced Lunch Status (FRL)		
	a. Qualifier for FRL	Binary
	b. Non-qualifier for FRL	Binary
c. Assessment Format		
	1. Pencil-and-Paper Based	Binary
	2. Computer-Based	Binary
Dependent Variables (Measurements)		
1. Algebra I EOC Score		Interval
2. 8th Grade Math MAP Score		Interval
3. ACT Mathematics Subsection Score		Interval

Grouping Factors

Students were filtered out by only including students that have a score for the
Algebra I EOC, 8th grade math MAP test, and an ACT score for mathematics to ensure
that mortality rate does not impact validity of the study (Creswell, 2009). From 2008
through 2009, some EOC assessments were given online as a pilot. District A then began
fully testing EOCs online in the 2009-2010 school year, C. Brinton (personal
communication, December 29, 2014). Therefore, only students that had taken the Algebra
I EOC during the 2009-2010 school year and later were selected and ended in students
that took the assessment on the 2011-2012 school year. Since the Algebra I EOC is
predominantly taken during the 9th grade year, then this timeframe would allow students

that tested in the 2011-2012 school year to have taken the ACT by the end of the 2014-2015 school year (their 12th grade year). By utilizing three years of student data, this study accounted for randomization of selection and maturation to help support internal validity (Creswell, 2009). The 8th grade MAP tests and ACT test were administered via P&P through this entire time period (Missouri DESE, 2014).

Data Collection

Archival data was identified to facilitate this study. Data was collected to answer the research questions posed in this study. Student test score data, free/reduced lunch status, and race data were necessary to complete the research. To determine which records were necessary to complete the research, permission was obtained from the University of Missouri Institutional Review Board (IRB) and the Director of Data Management for the School District A. Once the data was obtained, all individually identifiable information was removed from the data set by the data analyst.

The data sets that were used in this study were a comprehensive set of data identifiable by student number that was assigned by the school district. Within each data set for each student, there was an 8th grade math MAP test score, an Algebra I EOC test score, a math subset portion of the ACT score, the student's race, as well as their designation of whether the student qualified for free and reduced lunch or not.

Privacy and Security Measures

Privacy of the student data records was ensured. First, after all test scores were located and matched to the appropriate student, the data analyst removed all names. The data analyst removed all unique individual markers that could indicate the name or identity of a particular student. The IRB approval was obtained from the Chair of the IRB committee at the institution where the researcher was enrolled as a student. Additionally,

the researcher, prior to beginning any research, sought permission in the form of an official letter. This permission specified that the researcher was given the authority to utilize the data set provided by the Director of Data Management for School District A.

Data Analysis

Statistical analysis for this study was conducted using SPSS statistical analysis software and Microsoft Excel spreadsheet software. The intent of the analysis carried out for this study was to answer the five research questions outlined earlier in this chapter. Analyses were conducted including Pearson r Product Moment Coefficient of Correlation, ANCOVA, measures of centrality, effect size, and graphs and charts chronicling student performance on the ACT, Algebra I EOC, and 8th grade math MAP Assessments (Field, 2009). The information in Table 5 details the analysis strategy that was used to answer the research questions.

Table 5. *Summary of Analyses Used by Research Questions*

Research Question	Data Type	Analysis Strategy
Relationship between variables on 8th Grade MAP Assessment	Interval, Binary	Analysis of Covariance
Relationship between variables on Algebra I EOC	Interval, Binary	Analysis of Covariance
Relationship between variables on math subsection of ACT	Interval, Binary	Analysis of Covariance
Strength of Association	Interval, Binary	Analysis of Covariance

Summary

This chapter discussed the research methodology was divided into subsections. The statement of the problem and the purpose of the study was reviewed. Next, the research questions and hypothesis were provided. The population and sample were then presented. Finally, the research design, variables, grouping factors, data collection, privacy and security measures, and data analysis were clarified.

CHAPTER 4: PRESENTATION AND ANALYSIS OF DATA

Research shows that the results of prior studies regarding the use of technology to increase student achievement are mixed and contradictory (Attewell, 2001; Newhouse, 2011; Ricketts & Wilks, 2002). Studies have examined technology within education, but the studies largely examined the impact on instruction and less on the impact of assessment (Weglinsky, 1998). Therefore, further research is necessary to understand technology's effect on student outcome. The relationship not only needs to be examined in regards to the effect on student outcomes when technology is expanded in the instructional setting within education, but also when technology is incorporated more into assessments within education (Weglinsky, 1998).

Maguire, Smith, Brallier, and Palm (2008) conducted a study that compared assessment results for the same assessment administered by paper-and-pencil (P&P) versus computer-based using two different control groups. The study did not examine socio-economics or race (Maguire, et. al., 2010). Escudier, Newton, Cox, Reynolds, and Odell (2011) published a study that compared assessments results for dental undergraduate students that completed both paper-and-pencil and computer-based assessments (CBA) on the same content. Once again, the study did not examine socio-economics or race and was conducted in a higher educational setting (Escudier et al., 2011). Hargreaves, Shorrocks-Taylor, Swinnerton, Tait, and Threlfall (2010) conducted a similar study in England that had 10-year-old students take a P&P test in mathematics, and then a computer-based version of the test a month later. The authors argued that the computer-based test (CBT) had an overall positive effect on performance, but could not determine whether it was significant. Again, the study did not examine socio-economics or race and the students were taking a test that was not high-stakes (Hargreaves et. al.,

2010).

Kim and Huynh (2007) published a study that compared results of computer and P&P versions of both algebra and biology End of Course (EOC) assessments. In this study the authors argued that more research was needed on high-stakes assessments to investigate administration mode effects within the K-12 arena. Kim and Huynh looked at gender, race, and socio-economic status for the overall sample population, but did not examine whether any subgroups had different results then the norm. In this study, students took one version of the test and then soon after took another version of the same test. Kim and Huynh determined a limitation of the study was the voluntary participation. "Schools in which technology was more accessible to students might have been more likely to participate" (Kim and Huynh, 2007, p. 25). Kim and Huynh also argued that future studies should include samples with more diverse backgrounds due to the convenience nature of the sample. The authors admitted that African Americans and Hispanics were particularly underrepresented in the sample compare to the general makeup of the entire middle/high school students in the state. Also, with the requirements of the No Child Left Behind Act (NCLB) to provide assessment data on sub-group performance, Kim and Huynh argued that studies need to be done to look at mode effects for these subgroups specifically.

Poggio, Glasnapp, Yang, and Poggio (2005) examined score results for 7th grade students within mathematics testing in a large-scale assessment program. The study used the same assessment for both as the P&P assessment was transferred to computer. "Reasonable study controls support generalizability of these findings, but findings are limited to middle/level grade (7th grade) assessment of mathematics in a general education population" (Poggio, et. al., 2005, p. 26). This study was similar to Kim and

46

Huynh's (2007) study in which it was voluntary and students took both versions of the assessment. Specific research looking into public school K-12 student performance on existing assessments when technology is incorporated has been studied. However, the gap in research was still apparent when looking at high stakes assessments and considering socio-economics and race at the high school level. This study attempted to clear the present gap and help inform stakeholders to continue to make decisions about what format might be used to administer high stakes assessments moving forward.

Review of Research Design

The proposed research design utilized a static-group comparison design that utilizes a quantitative, non-experimental correlational methodology. Quantitative studies allow researchers to statistically analyze numerical data for reliability, significance, and correlation by using descriptive statistics (Creswell, 2009). The study used a static-group comparison design (Fraenkel & Wallen, 2003). A non-experimental correlational study was chosen because the independent variables were not manipulated at any point in the study (Creswell, 2009). This study did not attempt to establish a causal relationship; therefore, correlational methods were used to establish the degree to which the independent and dependent variables were related (Creswell, 2009).

The purpose of this study was to determine the extent to which P&P-based and computer-based high stakes assessments scores varied among a sample of 1836 students classified by race and socioeconomic status, located in School District A, in the state of Missouri, in 2014 and 2015. The P&P-based scores were for the 8th grade math Missouri Assessment Program (MAP) scale and the math portion of the ACT. The computer-based scores were for the Algebra I EOC scale. The student performance data were classified by two sub-groups (i.e., race, coded by 1 = minority group; 0 = not minority group, and

socio-economic status, coded by 1 = qualified for reduced/free lunch; 0 = did not qualify

for reduced/free lunch).

The results of the statistical analysis are presented systematically in five sections.

Each section presents the statistical evidence to address one research questions, as

follows: (a) What are the descriptive statistics for student performance on the 8th grade

math MAP test, the Algebra I EOC examinations, and the math portion of the ACT; (b) Is

there a difference in the student performance on the 8th grade math MAP examinations;

(c) Is there a difference in the student performance on the Algebra I EOC examinations;

(d) Is there a difference in student performance on the math portion of the ACT; and (e)

What is the difference in the strength of the association between the 8th grade math MAP

examination and the Algebra I EOC examination compared to the 8th grade math MAP

examination and the math portion of the ACT. Inferential statistics were used to test the

stated hypotheses.

Independent and Dependent Variables

Table 6 lists the independent and dependent variables explored in the study.

Table 6. *Independent and Dependent Variables Explored in the Study*

Variables
Independent Variables
1. Race
a.   Students From Minority Background
b.   Students Not From Minority Background
2. Free/Reduced Lunch Status (FRL)
a.   Qualifier for FRL
b.   Non-qualifier for FRL
3. Assessment Format
a.   Pencil-and-Paper Based
b.   Computer-Based
Dependent Variables
4. Algebra I EOC Score
5. 8th Grade Math MAP Score

Population

The selected research site, School District A, is located in the state of Missouri in the United States of America. The district is located in the suburbs of a major metropolitan area. C High School had their first students in the 2008-2009 school year with their first graduating class being in 2010.

Table 7. *Demographic Information*

School District A	2008	2009	2010	2011	2012	2013	2014
Total Enrollment	17,552	17,955	18,288	18,530	18,674	18,928	19,199
White Percent	71.7	70.5	69.7	65.9	64.8	64.3	63.6
Free/Reduced Lunch (FTE) Percent	38.7	40.9	43.8	46	47.6	48.7	49.6
School: A HIGH							
Total Enrollment	1,812	1,614	1,562	1,507	1,482	1,440	1,505
White Percent	63.5	60.5	56	53.3	50.9	49.2	49.6
Free/Reduced Lunch (FTE) Percent	40.8	43.4	46.5	50.6	52.9	55.2	55.8
School: B HIGH							
Total Enrollment	2122	1530	1314	1351	1382	1383	1413
White Percent	81.3	77.6	74.7	73.3	70	67.8	66.8
Free/Reduced Lunch (FTE) Percent	17.1	27.9	35.3	39.1	39.5	41.6	43.1
School: C HIGH							
Total Enrollment		947	1295	1347	1377	1377	1443
White Percent		80.6	80.4	79.8	78.6	77.7	76.2
Free/Reduced Lunch (FTE) Percent		15.9	16.9	17.4	18.7	17.7	19.1
School: D HIGH							
Total Enrollment	1636	1593	1516	1430	1346	1310	1325
White Percent	74.3	74	72.6	67.4	64.2	64.5	61.4
Free/Reduced Lunch (FTE) Percent	36.8	37.6	41.5	46.2	48	49.2	55.5

(DESE, mcds.dese.mo.gov, 2014)

In 2010 for example, there were 1389 graduates in the district with 57.88% taking the ACT (see Appendix B, ACT Results). This should then provide approximately 800 students for the sample size for that school year prior to pulling out students that did not have an 8th grade math MAP score and an Algebra I EOC score. The students were

selected from the four high schools from School District A because the district is one of the largest school districts in the state of Missouri, the district's student population represents a wide sampling of diversity and socio-economic status.

This school district was picked because the four high schools within the district spanned from urban to suburban locations and had a wide variance of free and reduced and white versus non-white. The study was limited to the four high schools within this district because all four high schools were exposed to 1:1 at the same time and with the same devices, the schools had the same curriculum guides, and testing plans. All of the high school students within School District A were given a laptop computer that the students used and took home with them during the school year for all four years of high school.

To ensure that only student data that was considered was for students that there were complete data sets, the student results were then narrowed down to only students that had taken all three of the following: 8th grade mathematics MAP, Algebra I EOC in Missouri, and the ACT. As a result, the data included 1,836 students from the HS classes of 2013, 2014, and 2015 that met the aforementioned criteria.

Statistical Analysis

The study used an Analysis of covariance (ANCOVA) to evaluate difference among group means. Including covariates in Analysis of Variance (ANOVA) helps remove any bias of variables that may be unmeasured and it reduces the error variance (Field, 2009). This allowed for an analysis of data utilizing both subsections of data gathered (i.e. Caucasian, Race, etc.). McDonald (2014) suggested using ANCOVA to compare two or more regression lines to each other, which helps to determine if the lines are different in either slope or intercept.

50

P-Value

The researcher established the *p*-value at .05 and rejected the null hypothesis

when the *p*-value was smaller than or equal to 0.05. If the *p*-value was rejected, then the

researcher used post-hoc tests. The analysis utilized Pearson's-*r* product moment

coefficient of correlation, ANCOVA, measures of centrality, standard deviations, graphs,

figures, and scatter plots to present the findings and facilitate the analysis of the results

(Field, 2009).

Null Hypotheses

In order to investigate the problem, address the purpose, and to answer the

research questions in this study, the following null hypotheses were tested.

$H_0$1. There is no difference in descriptive statistics of student performance on the

8th grade math MAP test, the Algebra I EOC examinations, and the math portion of the

ACT when considering race (minority and not minority) and free/reduced lunch status

(non-qualifier and qualifier).

$H_0$2. There is no difference in the relationship between student performance on

the 8th grade math MAP examinations when considering race (minority and not minority)

and free/reduced lunch status (non-qualifier and qualifier).

$H_0$3. There is no difference in the relationship between student performance on

the Algebra I EOC examinations when considering race (minority and not minority) and

free/reduced lunch status (non-qualifier and qualifier).

$H_0$4. There is no difference in the relationship between student performance on

the Algebra I EOC examinations when considering race (minority and not minority) and

free/reduced lunch status (non-qualifier and qualifier).

$H_0$5. There is no difference in the measurement of strength of association between

the 8th grade math MAP examination and the Algebra I EOC examination compared to the 8th grade math MAP examination and the math portion of the ACT when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

Research Question 1: Descriptive Statistics for Student Performance

The scores for the 8th grade math MAP scale ranged from 525 to 885. The scores for the math portion of the ACT ranged from 12 to 36. The scores for the Algebra I EOC scale ranged from 151 to 250. Consequently, the descriptive statistics for the three types of assessment were not directly comparable. The scores for each type of assessment were transformed to percentages of the maximum score, in order that the descriptive statistics could be directly compared. The descriptive statistics for the percentage scores are compared in Table 8.

Table 8. *Descriptive Statistics for Student Performance*

Statistic	ACT Math Paper and Pencil (%)	MAP Math Scale Paper and Pencil (%)	Algebra I EOC Scale Computer-based (%)
M	57.67	81.62	84.48
Mdn	55.56	81.58	84.00
Mode	47.22	81.58	82.00
SD	12.65	3.75	6.84
Skew	0.63	-0.05	-0.04
Min	33.33	59.32	60.40
Max	100.00	100.00	100.00
Range	66.67	40.68	39.60

The frequency distributions of the scores are illustrated in Figure 3. The MAP math scale and Algebra I EOC scale were normally distributed, reflected by approximation to bell- curves, low skewness (Skew – 0.05 and – 0.04 respectively); and central tendency (*Mode* = 81.58 and 82.00, with modes close to the mean scores, $M =$ 81.62 and 84.48 respectively). In contrast, the scores for the math portion of the ACT

were positively skewed, reflected by deviation from a bell-shaped curve, high positive skewness (*Skew* = 0.63) and a non-central mode (*Mode* = 47.72) less than the mean (M = 57.67). The variability in the ACT math scores (Range = 66.67, SD = 12.65) was greater than the variability in the MAP math scale (*Range* = 40.69, *SD* = 3.75) and the Algebra I EOC scale (*Range* = 39.60, *SD* = 6.84).
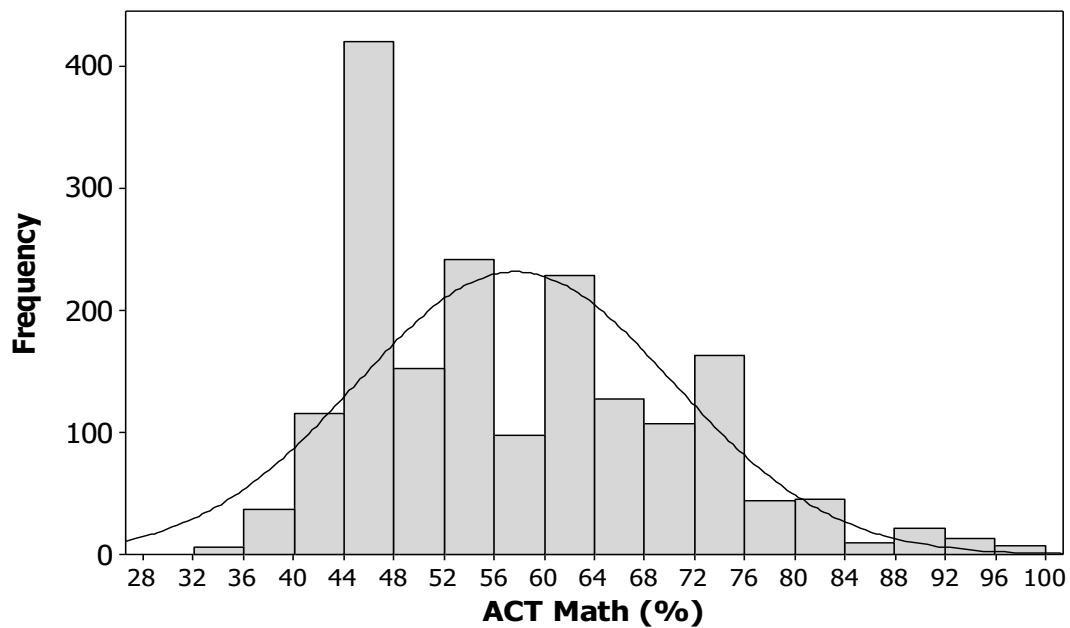
Figure 3. Frequency distribution histogram of the scores for the Math portion of the ACT

N=183

Figure 4. Frequency distribution histogram of the scores for the Algebra I EOC N=1836
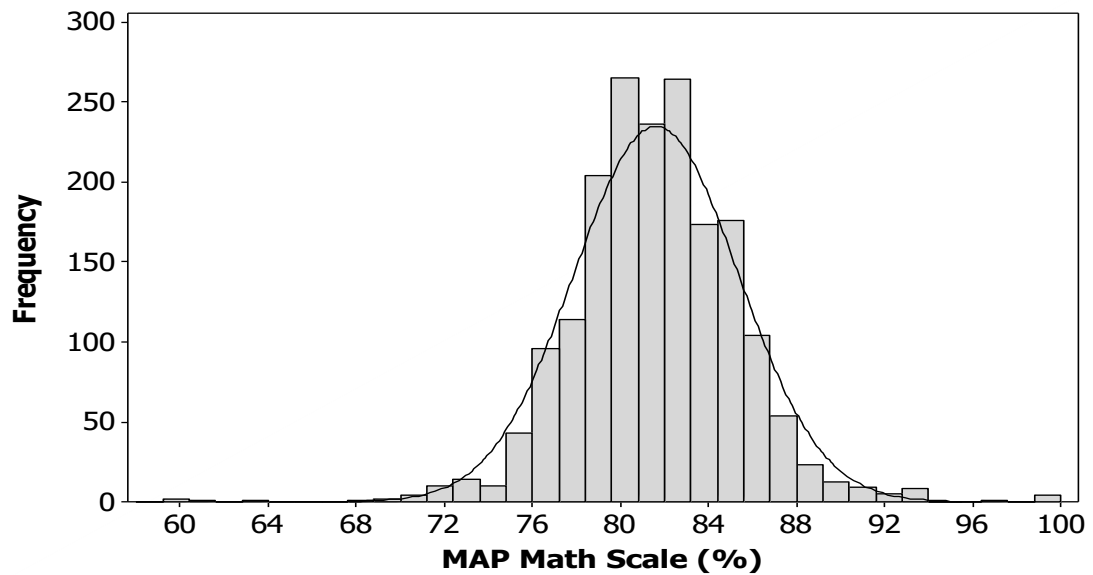
Figure 5. Frequency distribution histogram of the scores for the 8th grade Math MAP assessment N=1836

Table 9 presents the results of a one-way analysis of variance to compare the mean scores for the three types of assessment.

Table 9. *One-way Analysis of Variance (ANOVA) on Score for Three Types of Assessments*

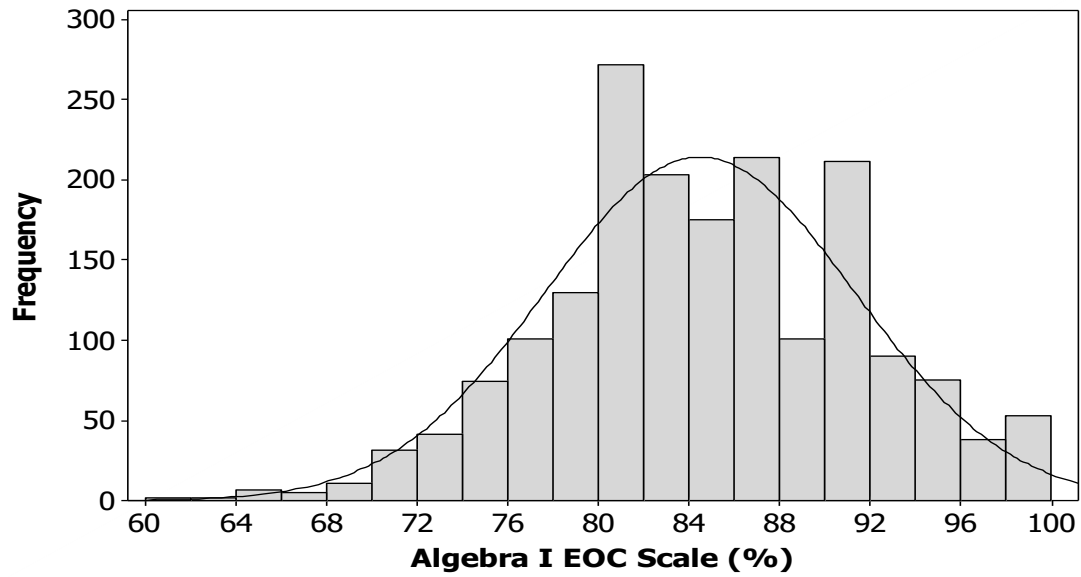| Source	Type III SS	df	MS	F	p	Partial ETA Squared
Type of Assessment	795616.24	2	397808.12	5404.73	<.001*	.663
Error	405188.35	5505	73.60			
Total	31845577.90	5508				
Corrected Total	1200804.59	5507				

* Significant (p < .001)

Statistical evidence was presented to reject H_01: There is no difference in descriptive statistics of student performance on the 8th grade math MAP test, the Algebra I EOC examinations, and the math portion of the ACT. The scores for ACT math ($M = 57.67, SD = 12.65$) were significantly lower ($F (2, 5505) = 5404.73, p < .001$) than the scores for both the MAP math scale ($M = 81.62, SD = 3.75$) and the Algebra I EOC scale ($M = 84.48, SD = 6.84$). The effect size, representing the proportion of the variance explained, was large (Partial Eta Squared = .663). It cannot, however, be concluded that there is a difference between paper and pencil based scores (for the 8th grade math MAP scale) and the computer-based scores (for the Algebra I EOC scale) because there was no significant difference between the mean scores for these two measures of student performance.

Research Question 2: Student Performance on the MAP Math Scale

The descriptive statistics of the scores for the MAP math scale, classified by Socio Economic Status (SES) and Race are summarized in Table 10. The results of factorial ANOVA, with the scores as the dependent variable, and with SES and race as the two fixed factors, assuming equality of variance, are presented in Table 11.

Table 10. *Descriptive Statistics of the Scores for the MAP Math Scale*

SES	Race	M	SD	N
No Free/Reduced Lunch	Not Minority	82.52	3.58	1010
	Minority	81.00	3.54	286
	Total	82.18	3.62	1296
Free/Reduced Lunch	Not Minority	81.23	3.42	231
	Minority	79.53	3.73	309

	Total	80.26	3.69	540
Total	Not Minority	82.28	3.59	1241
	Minority	80.24	3.71	595
	Total	81.62	3.75	1836

Table 11. *Factorial Analysis of Variance (ANOVA) on the Scores for the MAP Math Scale*

Source	Type III SS	df	MS	F	p	Partial ETA Squared
SES	630.25	1	630.25	49.20	<.001*	.026
Race	855.52	1	855.52	66.79	<.001*	.035
SES* Race	2.723	1	2.72	0.213	.645	.000
Error	23467.59	1832	12.81			
Total	12256146.63	1836				
Corrected Total	25769.893	1835				

Note: * Significant (p < .001)

Statistical evidence was presented to reject $H_0 2$: There is no difference in the relationship between student performance on the 8th grade math MAP examinations when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier). The scores for the minority group ($M = 80.24$, $SD = 3.71$) were significantly lower ($F (1, 1832) = 66.79$, $p < .001$) than the scores for the not minority group ($M = 82.28$, $SD = 3.59$). The scores for students who qualified for free/reduced lunch ($M = 80.26$, $SD = 3.69$) were significantly lower ($F (1, 1832) = 49.20$, $p < .001$) than for the students who did not qualify ($M = 82.18$, $SD = 3.58$). The effect sizes, however (Partial Eta Squared = .026 and .035) were very small. There was no significant interaction between SES and Race.

Research Question 3: Student Performance on the Algebra I EOC Scale

The descriptive statistics of the scores for the Algebra I EOC scale, classified by SES and Race are summarized in Table 12. The results of factorial ANOVA, with the scores as the dependent variable, and with SES and Race as the two fixed factors, assuming equality of variance, are presented in Table 13.

56

Table 12. *Descriptive Statistics of the Scores for the Algebra I EOC Scale*

SES	Race	M	SD	N
No Free/Reduced Lunch	Not Minority	86.14	6.41	1010
	Minority	83.35	6.65	286
	Total	85.53	6.57	1296
Free/Reduced Lunch	Not Minority	83.60	6.88	231
	Minority	80.75	6.57	309
	Total	81.97	6.85	540
Total	Not Minority	85.67	6.58	1241
	Minority	82.00	6.73	595
	Total	84.48	6.84	1836

Table 13. *Factorial Analysis of Variance (ANOVA) on the Scores for the Algebra I EOC Scale*

Source	Type III SS	df	MS	F	p	Partial ETA Squared
SES	2193.12	1	2193.12	51.32	<.001*	0.027
Race	2641.10	1	2641.10	61.80	<.001*	0.033
SES* Race	0.29	1	0.29	0.01	0.934	0.000
Error	78296.25	1832	42.74			
Total	13188937.44	1836				
Corrected Total	85927.24	1835				

Note: * Significant (p < .05)

Statistical evidence was presented to reject $H_0 3$: There is no difference in the relationship between student performance on the Algebra I EOC examinations when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier). The scores for the minority group (*M* = 82.00, *SD* = 6.73) were significantly lower (*F* (1, 1832) = 61.80, *p* < .001) than the scores for the not minority group (*M* = 85.67, *SD* = 6.58). The scores for students who qualified for free/reduced lunch (*M* = 81.97, *SD* = 6.85) were significantly lower (*F* (1, 1832) = 51.32, *p* < .001) than for the students who did not qualify (*M* = 85.53, *SD* = 3.58). The effect sizes, however (Partial Eta Squared = .027 and .033) were very small. There was no significant interaction between SES and Race.

Research Question 4: Student Performance on the Math Portion of the ACT

The descriptive statistics of the scores for the math portion of the ACT, classified by SES and Race are summarized in Table 14. The results of factorial ANOVA, with the scores as the dependent variable, and with SES and Race as the two fixed factors, assuming equality of variance, are presented in Table 15.

Table 14. *Descriptive Statistics of the Scores for the Math Portion of the ACT*

SES	Race	M	SD	N
No Free/Reduced Lunch	Not Minority	60.84	12.94	1010
	Minority	55.73	11.64	286
	Total	59.71	12.83	1296
Free/Reduced Lunch	Not Minority	55.25	11.24	231
	Minority	50.94	9.93	309
	Total	52.79	10.72	540
Total	Not Minority	59.80	12.82	1241
	Minority	53.24	11.04	595
	Total	57.67	12.65	1836

Table 15. *Factorial Analysis of Variance (ANOVA) on the Scores for the Math Portion of the ACT*

Source	Type III SS	df	MS	F	p	Partial Eta Squared
SES	8918.77	1	8918.77	61.20	<.001*	.032
Race	7357.77	1	7357.77	50.49	<.001*	.027
SES* Race	52.40	1	52.40	0.36	0.549	.000
Error	266964.36	1832	145.72			
Total	6400493.83	1836				
Corrected Total	293491.22	1835				

Note: * Significant (p < .05)

Statistical evidence was presented to reject H_04: There is no difference in the relationship between student performance on the math portion of the ACT when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier). The scores for the minority group ($M = 53.24$, $SD = 11.04$) were significantly lower ($F (1, 1832) = 50.49$, $p < .001$) than the scores for the not minority group ($M = 59.80$, $SD = 12.82$). The scores for students who qualified for free/reduced lunch ($M = 52.79$, $SD = 10.72$) were significantly lower ($F (1, 1832) = 61.20$, $p < .001$)

than for the students who did not qualify ($M = 59.71$, $SD = 12.83$). The effect sizes, however (Partial Eta Squared = .032 and .027), were very small. There was no significant interaction between SES and Race.

Research Question 5: Strength of the Association between the Assessments

Figure 6 presents a matrix plot of the scores for the Algebra I EOC scale, the 8th grade math MAP scale, and the math portion of the ACT. The matrix plot visually indicates that the scores for the three types of assessment were positively correlated with each other.
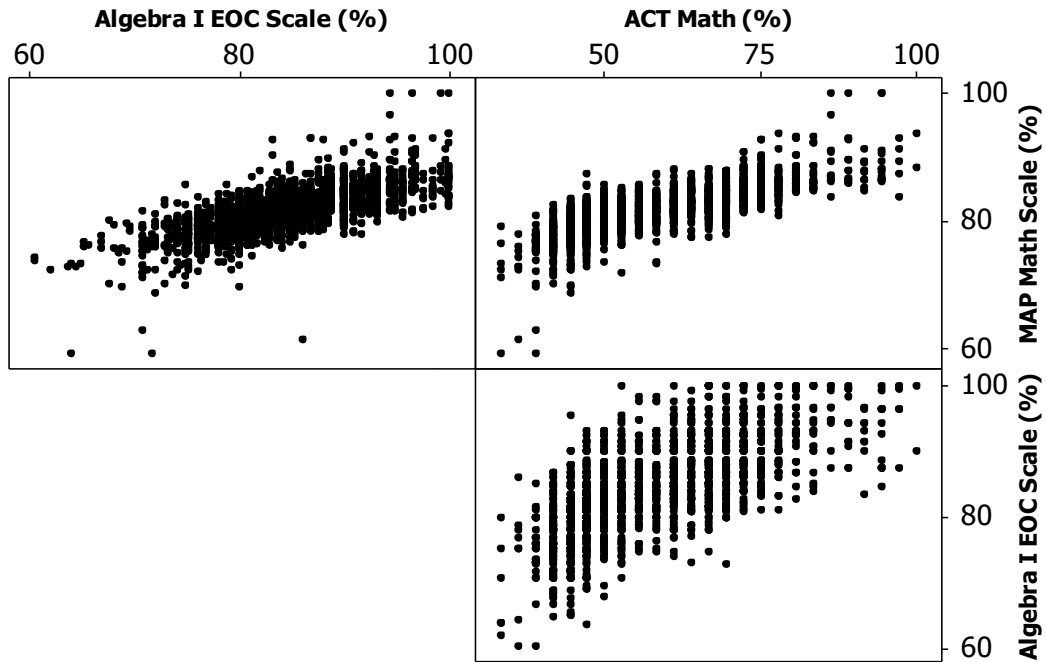
Figure 6. Matrix plot of the scores for the three types of assessments

The matrix of Pearson's correlation coefficients in Table 16 confirms that the scores for the three types of assessment were significantly and positively correlated. The strongest correlation was between the MAP math scale and ACT math (Pearson's r ($N = 1836$) = .715, $p < .001$). The weakest correlation was between the Algebra I EOC scale

and ACT math (Pearson's r (N = 1836) = .683, p < .001).

Table 16. *Correlation Matrix between Scores for the Three Types of Assessments*

	ACT Math	Algebra I EOC Scale	MAP Math Scale
ACT Math	1		
Algebra I EOC Scale	.683*	1	
MAP Math Scale	.778*	.715*	1

Note: * Significant (p < .001)

The linear relationship between the ACT math vs. Algebra I EOC scores, stratified by SES, is illustrated in Figure 7, and stratified by Race in Figure 8. The fitted linear regression lines in Figure 7 and 8 appeared to cross, suggesting a disordinal interaction, and reflecting a significant difference between the slopes of the linear regression lines with respect to SES and Race.

Figure 7. Relationship between ACT Math and Algebra I EOC Scores, Stratified by SES
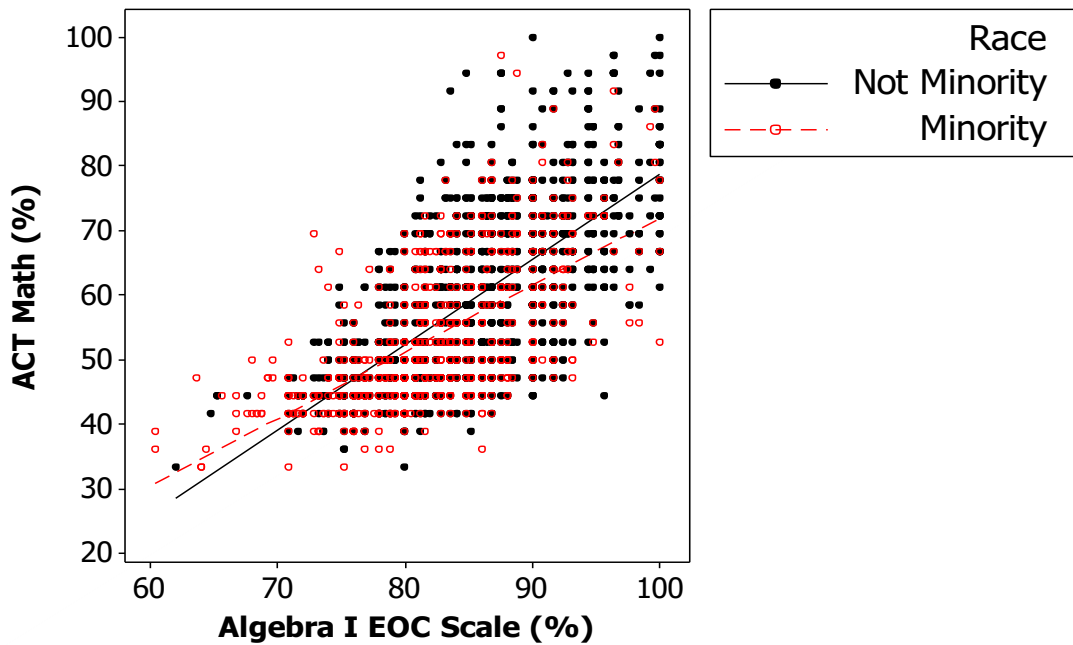
Figure 8. Relationship between ACT Math and Algebra I EOC scores, stratified by Race

Table 17 confirmed the significance of this interaction using ANCOVA.

Table 17. *Analysis of covariance (ANCOVA) on Interaction of SES and Race between ACT Math versus Algebra I EOC Scores*

Source	Type III SS	df	MS	F	p	Partial ETA Squared
SES * Algebra I EOC	985.79	1	985.79	11.94	.001*	.006
Race * Algebra I EOC	739.02	1	739.02	8.95	.003*	.005
Error	151090.38	1830	82.56			
Total	6400493.83	1836				
Corrected Total	293491.22	1835				

Note: * Significant (p < .01)

The interaction with SES was significant ($F_{(1, 1830)}$ = 11.94, p = .001) and the interaction with Race was also significant ($F_{(1, 1830)}$ = 8.95, p = .003). Consequently, the slopes of the regression lines illustrated in Figures 5 and 6 were significantly different. The effect sizes, however, were very low (.006 and .005 respectively). Statistically, a large sample size yields significance, but these findings do not suggest a sizeable effect. Statistical evidence was provided to reject one part of $H_0 5$: There is no difference in the measurement of strength of association between the Algebra I EOC

examination compared to the math portion of the ACT when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

The linear relationship between the ACT math vs. MAP math scores, stratified by SES, is illustrated in Figure 9, and stratified by Race in Figure 10. The linear relationship between the MAP math and Algebra I EOC scores, stratified by SES, is illustrated in Figure 9, and stratified by Race in Figure 10. The fitted linear regression lines in Figure 11 and 12 appeared to cross, suggesting a disordinal interaction, and reflecting a significant difference between the slopes of the linear regression lines with respect to SES and Race. Table 18 confirmed the significance of this interaction using ANCOVA.
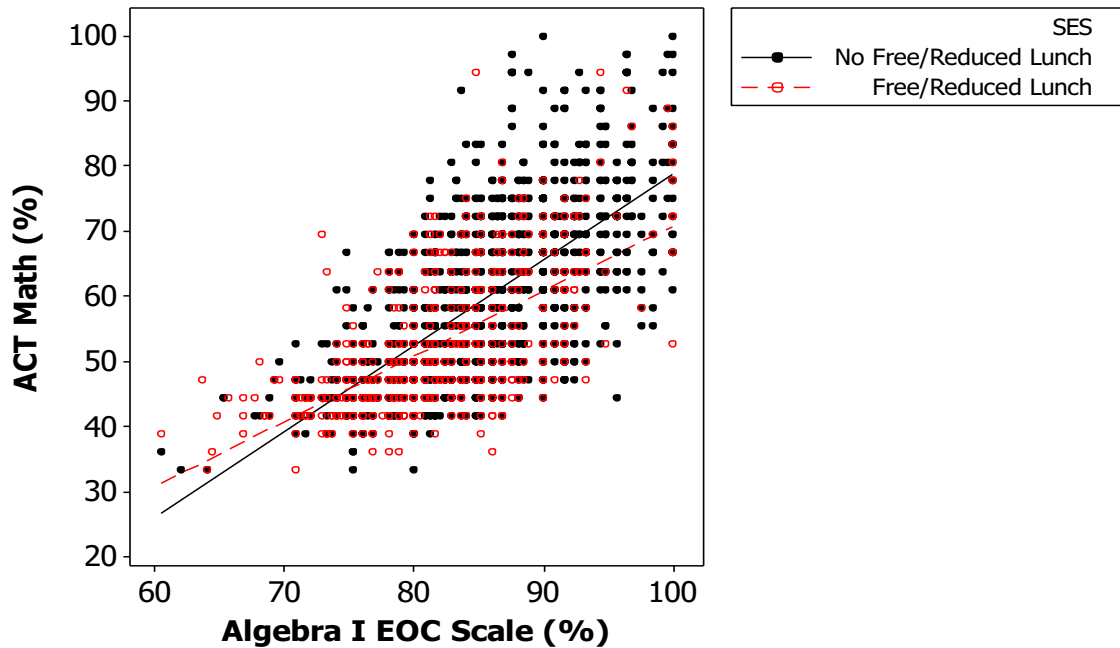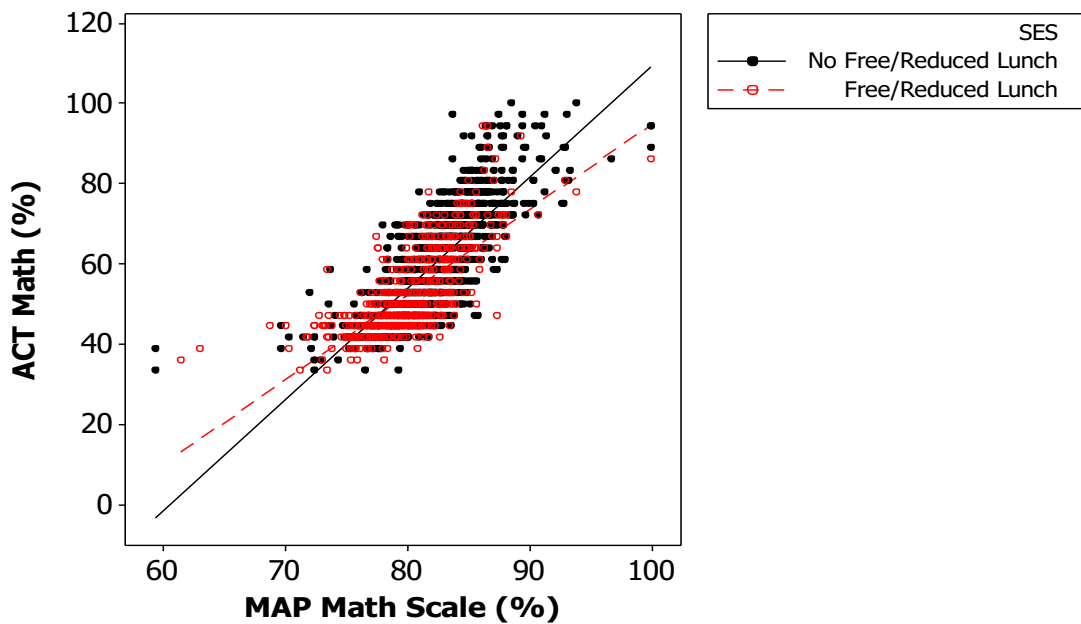

Figure 9. Relationship between ACT Math and MAP Math scores, stratified by SES
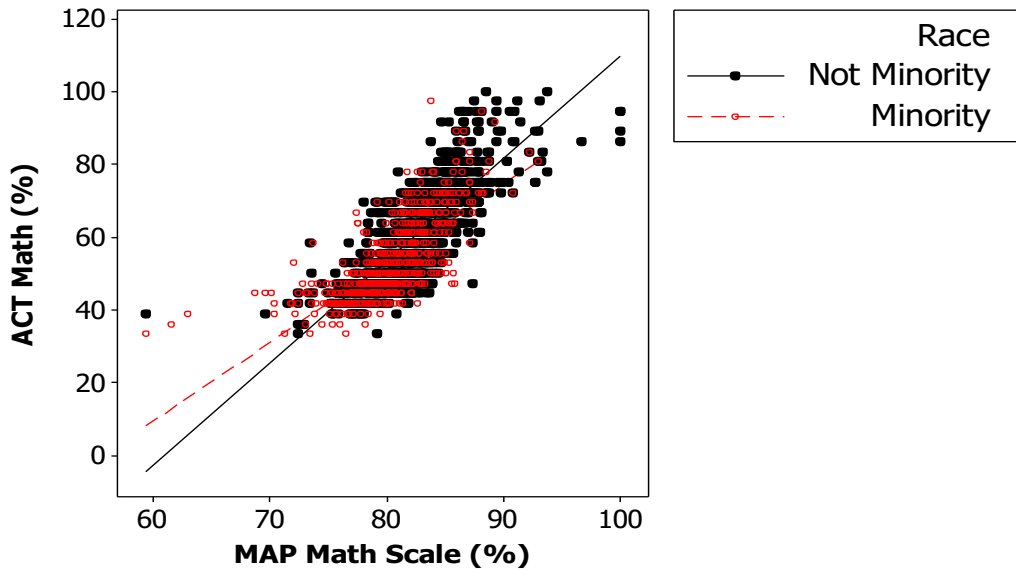
Figure 10. Relationship between ACT Math and MAP Math scores, stratified by Race

Table 18. *Analysis of covariance (ANCOVA) on Interaction of SES and Race between ACT Math versus MAP Math Scores*

Source	Type III SS	*df*	*MS*	*F*	*p*	Partial ETA Squared
SES * MAP Math	876.52	1	876.52	14.47	<.001*	0.008
Race * MAP Math	1048.10	1	1048.10	17.30	<.001*	0.009
Error	110865.16	1830	60.58			
Total	6400493.83	1836				
Corrected Total	293491.22	1835				

Note: * Significant (p < .001)

The interaction with SES was significant ($F (1, 1830) = 14.47, p < .001$) and the interaction with Race was also significant ($F (1, 1830) = 17.30, p < .001$). Consequently, the slopes of the regression lines illustrated in Figures 7 and 8 were significantly different. The effect sizes, however, were very low (.008 and .009 respectively). Statistically, a large sample size yields significance, but these findings do not suggest a sizeable effect. Statistical evidence was provided to reject one part of $H_0 5$: There is no difference in the measurement of strength of association between the 8th grade MAP math examination compared to the math portion of the ACT when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

The linear relationship between the MAP math versus Algebra I EOC scores, stratified by SES, is illustrated in Figure 11, and stratified by Race in Figure 12. The fitted linear regression lines in Figure 11 and 12 appeared to be parallel, suggesting no interaction, and reflecting no significant difference between the slopes of the linear regression lines with respect to SES and Race.
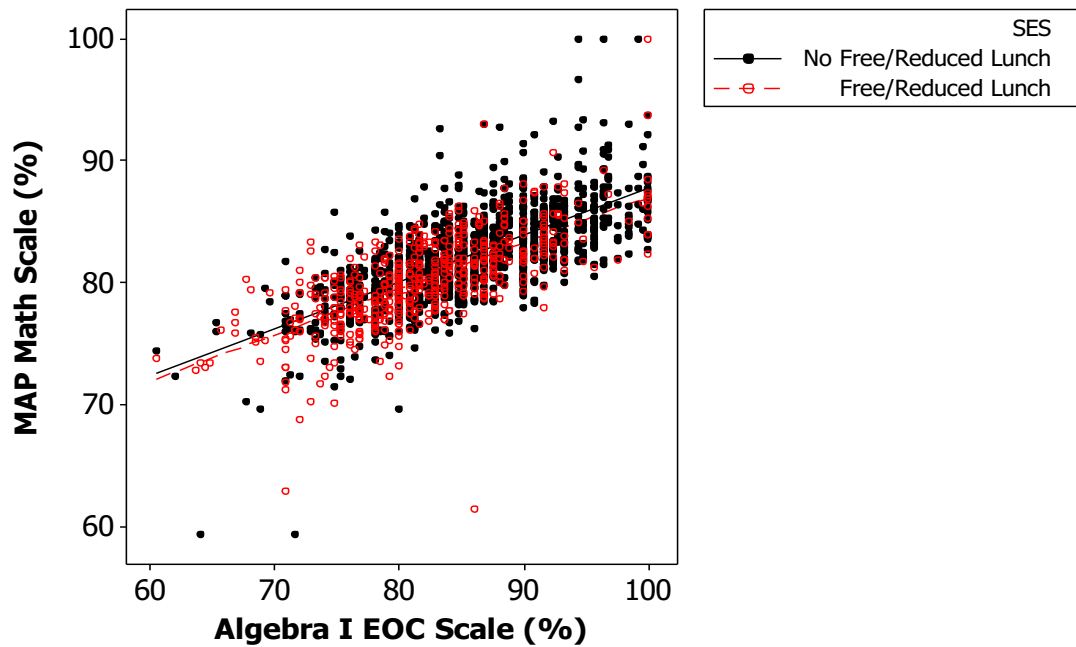
Figure 11. Relationship between MAP Math and Algebra I EOC scores, stratified by SES
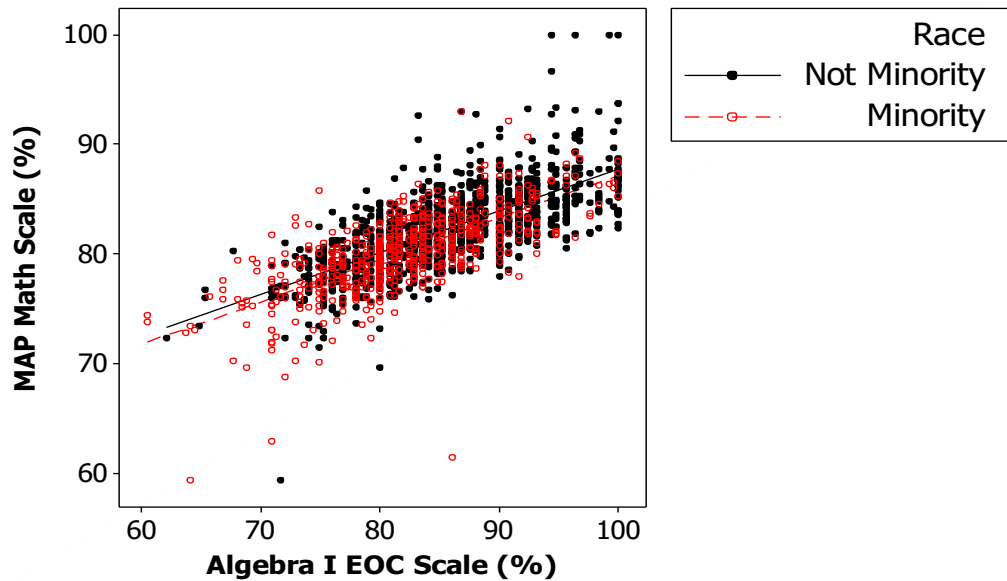
Figure 12. Relationship between MAP Math and Algebra I EOC scores, stratified by Race

Table 19 confirmed the lack of significance of this interaction using ANCOVA.

Table 19. *ANCOVA on Interaction of SES and Race between MAP Math versus Algebra I EOC Scores*

Source	Type III SS	df	MS	F	p	Partial ETA Squared
SES * Algebra I EOC	2.25	1	2.25	0.33	.564	.000
Race * Algebra I EOC	0.91	1	0.91	0.13	.714	.000
Error	12371.36	1830	6.76			
Total	12256146.64	1836				
Corrected Total	25769.89	1835				

Note: * Significant (p < .001)

The interactions with SES and Race were not significant ($p > .05$) and the effect sizes were zero. Consequently, the slopes of the linear regression lines illustrated in Figures 11 and 12 were not significantly different. There was insufficient statistical evidence to reject the final part of $H_0 5$: There is no difference in the measurement of strength of association between the 8th grade MAP math examination compared to Algebra I EOC examination when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

Overall Findings

After the scores for each type of assessment were transformed to percentages of the maximum score, in order that the descriptive statistics could be directly compared, the scores for the math portion of the ACT were found to be significantly lower than the scores for both the MAP math scale and the Algebra I EOC scale. A conclusion cannot be drawn that there is a difference between paper and pencil based scores (for the 8th grade math MAP scale) and the computer-based scores (for the Algebra I EOC scale) because there was no significant difference between the mean scores for these two measures of student performance.

For all three assessments analyzed, the 8th grade math MAP, the Algebra I EOC, and the math portion of the ACT, the scores for the minority group were significantly lower than the scores for the not minority group. The scores for students who qualified for free/reduced lunch were also significantly lower than for the students who did not qualify. The effect sizes, however were very small. Overall, where there was a difference in scores for students that qualified as minority or for free and reduced lunch status from students that did not qualify, there was no significant impact on student assessments results based on paper-and-pencil compared to computer-based assessments according to the data analyzed.

Chapter 5 provides suggestions for future research. This future research would further contribute to the general body of knowledge while potentially allowing stakeholders that make decisions on testing formats and models to make more informed decisions. Chapter 5 also lists the findings, conclusions, and limitations of the study.

CHAPTER 5: FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

This chapter provides an overview of the problem and purpose of this study, population, and statistical methods, as well as a discussion of the findings, conclusions, recommendations, and limitations of the study.

Problem of the Study

There is a lack of knowledge in whether there is a difference in results for students on paper and pencil high stakes assessments and computer-based high stakes assessments when considering race and/or free and reduced lunch status. Herold (2016, para. 1) asserted that "In 2015-16, for the first time, more state standardized tests for the elementary and middle grades will be administered via technology than by paper and pencil." Herold went on to state that both academics and parents alike have concerns that achievement gaps may be widened as student have unequal access to and use of technology. It is a reasonable belief to think that the increase of technologies use within education will help to level the playing field for current racial, gender, and geographic divides (Noeth, 2004). This cannot be made as an absolute, however, without data to support the belief. Is there a difference in results for students on paper-and-pencil (P&P) high stakes assessments and computer-based high stakes assessments when considering free and reduced lunch status and/or race? This study looked to close the knowledge gap in the research and examine whether there is an increased effect on the difference between the 8th grade Missouri Assessment Program (MAP) assessment and the Algebra I End of Course (EOC) test results in comparison to the difference between the 8th grade MAP assessment and ACT test results when taking into consideration race and socio-economics of the student.

Purpose of the Study

67

The purpose of this study was to add new knowledge to this field of study by determining whether there is a difference in results for students on P&P high stakes assessments and computer-based high stakes assessments when considering free and reduced lunch status and/or race. To accomplish this task, this study took into consideration the established relationships of a P&P assessment and the use of an online assessment. By contributing further to the existing research, the findings of this study may provide educators and policy makers additional information to make decisions that inform future policy discussions regarding the use of computer-based assessments for high stakes assessments within the K-12 public education setting.

Population

The selected research site, School District A, is located in the state of Missouri in the United States of America. The district is located in the suburbs of a major metropolitan area. C High School had their first students in the 2008-2009 school year with their first graduating class in 2010. In 2010 for example, there were 1,389 graduates in the district, of which 57.88% took the ACT. This data provided approximately 800 students for the sample size for that school year prior to pulling out students that did not have an 8th grade math MAP score and an Algebra I EOC score. The students were selected from the four high schools from School District A because the district is one of the largest school districts in the state of Missouri, the district's student population represents a wide sampling of diversity and socio-economic status.

This school district was picked because the four high schools within the district spanned from urban to suburban locations and had a wide variance of free and reduced and white versus non-white. The study was limited to the four high schools within this district because all four high schools were exposed to 1:1 at the same time and with the

68

same devices, the schools had the same curriculum guides, and testing plans. All of the high school students within School District A were given a laptop computer that the students used and took home with them during the school year for all four years of high school. To ensure that only student data that was considered was for students that there were complete data sets, the student results were then narrowed down to only students that had taken all three of the following: 8th grade mathematics MAP, Algebra I EOC in Missouri, and the ACT. As a result, the data included 1,836 students from the HS classes of 2013, 2014, and 2015 that met the aforementioned criteria.

Statistical Methods

The study used an Analysis of covariance (ANCOVA) to evaluate difference among group means. Including covariates in Analysis of Variance (ANOVA) helps remove any bias of variables that may be unmeasured and it reduces the error variance (Field, 2009). This allowed for an analysis of data utilizing both subsections of data gathered (i.e. Caucasian, Race, etc.). McDonald (2014) suggested using ANCOVA to compare two or more regression lines to each other, which helps to determine if the lines are different in either slope or intercept.

The researcher established the *p*-value at .05 and rejected the null hypothesis when the *p*-value was smaller than or equal to .05. If the *p*-value was rejected, then the researcher pursued post-hoc tests and the results were displayed and analyzed. The analysis utilized Pearson's-*r* product moment coefficient of correlation, ANCOVA, measures of centrality, standard deviations, graphs, figures, and scatter plots to present the findings and facilitate the analysis of the results (Field, 2009).

Discussion of Study Findings

The discussion of findings is organized by research question. Careful examination

of the data collected in order to answer the problem and research questions in this study led to the following findings.

Research Question 1: What are the descriptive statistics for student performance on the 8th grade math MAP test, the Algebra I EOC examinations, and the math portion of the ACT broken out by the following subgroups?

By Race:

    a. Minority

    b. Not Minority

By Free/Reduced Lunch Status:

    c. Qualifiers

    d. Non-qualifiers

$H_0 1$. There is no difference in descriptive statistics of student performance on the 8th grade math MAP test, the Algebra I EOC examinations, and the math portion of the ACT when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

Descriptive statistics calculated for the independent variables revealed:

1. The scores for the 8th grade math MAP scale ranged from 12 to 36. The scores for the math portion of the ACT ranged from 525 to 885. The scores for the Algebra I EOC scale ranged from 151 to 250.

2. The MAP math scale and Algebra I EOC scale were normally distributed, reflected by approximation to bell curves, low skewness; and central tendency. In contrast, the scores for the math portion of the ACT were positively skewed, reflected by deviation from a bell-shaped curve, high positive skewness, and a non-central mode less than the mean.

3. The variability in the ACT math scores was greater than the variability in the MAP math scale and the Algebra I EOC scale.

Statistical evidence was presented to reject $H_0 1$: There is no difference in descriptive statistics of student performance on the 8th grade math MAP test, the Algebra I EOC examinations, and the math portion of the ACT. The scores for ACT math were significantly lower than the scores for both the MAP math scale and the Algebra I EOC scale. The effect size, representing the proportion of the variance explained, was large. It cannot, however, be concluded that there is a difference between P&P-based scores (for the 8th grade math MAP scale) and the computer-based scores (for the Algebra I EOC scale) because there was no significant difference between the mean scores for these two measures of student performance.

Research Question 2: Is there a difference in the relationship between student performance on the 8th grade math MAP examinations when broken out by the following subgroups?

By Race:

    a. Minority

    b. Not Minority

By Free/Reduced Lunch Status:

    c. Qualifiers

    d. Non-qualifiers

$H_0 2$. There is no difference in the relationship between student performance on the 8th grade math MAP examinations when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

The descriptive statistics of the scores for the MAP math scale, classified by

Socio Economic Status (SES) and Race, were analyzed. The results of factorial ANOVA, with the scores as the dependent variable, and with SES and Race as the two fixed factors, assuming equality of variance was also analyzed. Statistical evidence was presented to reject $H_0 2$: There is no difference in the relationship between student performance on the 8th grade math MAP examinations when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier). The scores for the minority group were significantly lower than the scores for the not minority group. The scores for students who qualified for free/reduced lunch were significantly lower than for the students who did not qualify. The effect sizes, however were very small. There was no significant interaction between SES and Race.

Research Question 3: Is there a difference in the relationship between student performance on the Algebra I EOC examinations when broken out by the following subgroups?

By Race:

    a. Minority

    b. Not Minority

By Free/Reduced Lunch Status:

    a. Qualifiers

    b. Non-qualifiers

$H_0 3$. There is no difference in the relationship between student performance on the Algebra I EOC examinations when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

The descriptive statistics were analyzed of the scores for the Algebra I EOC scale, classified by SES and Race. The results of factorial ANOVA, with the scores as the

dependent variable, and with SES and Race as the two fixed factors, assuming equality of

variance, were also determined. Statistical evidence was presented to reject H_03: There is

no difference in the relationship between student performance on the Algebra I EOC

examinations when considering race (minority and not minority) and free/reduced lunch

status (non-qualifier and qualifier). The scores for the minority group were significantly

lower than the scores for the not minority group. The scores for students who qualified

for free/reduced lunch were significantly lower than for the students who did not qualify.

The effect sizes, however were very small. There was no significant interaction between

SES and Race.

Research Question 4: Is there a difference in the relationship between student

performance on the math portion of the ACT when broken out by the following

subgroups?

By Race:

a. Minority

b. Not Minority

By Free/Reduced Lunch Status:

a. Qualifiers

b. Non-qualifiers

H_04. There is no difference in the relationship between student performance on

the Algebra I EOC examinations when considering race (minority and not minority) and

free/reduced lunch status (non-qualifier and qualifier).

The descriptive statistics of the scores for the math portion of the ACT, classified

by SES and Race were analyzed. The results of factorial ANOVA, with the scores as the

dependent variable, and with SES and Race as the two fixed factors, assuming equality of

variance, were also analyzed. Statistical evidence was presented to reject $H_0 4$: There is no difference in the relationship between student performance on the math portion of the ACT when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier). The scores for the minority group were significantly lower than the scores for the not minority group. The scores for students who qualified for free/reduced lunch were significantly lower than for the students who did not qualify. The effect sizes, however were very small. There was no significant interaction between SES and Race.

Research Question 5: What is the difference in the measurement of strength of association between the 8th grade math MAP examination and the Algebra I EOC examination compared to the 8th grade math MAP examination and the math portion of the ACT when broken out by the following subgroups?

By Race:

    c. Minority

    d. Not Minority

By Free/Reduced Lunch Status:

    c. Qualifiers

    d. Non-qualifiers

$H_0 5$. There is no difference in the measurement of strength of association between the 8th grade math MAP examination and the Algebra I EOC examination compared to the 8th grade math MAP examination and the math portion of the ACT when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

A matrix plot analysis determined that the scores for the three types of assessment were positively correlated with each other. The matrix of Pearson's correlation

coefficients confirmed that the scores for the three types of assessment were significantly and positively correlated. The strongest correlation was between the MAP math scale and ACT math. The weakest correlation was between the Algebra I EOC scale and ACT math.

The linear relationship between the ACT math versus Algebra I EOC scores, stratified by SES, and stratified by Race appeared to have crossing fitted linear regression lines, suggesting a disordinal interaction, and reflecting a significant difference between the slopes of the linear regression lines with respect to SES and Race. Further analysis confirmed the significance of this interaction using ANCOVA.

The interaction with SES was significant and the interaction with Race was also significant. Consequently, the slopes of the regression lines for SES and Race respectively were significantly different. The effect sizes, however, were very low. Statistical evidence was provided to reject one part of H_05: There is no difference in the measurement of strength of association between the Algebra I EOC examination compared to the math portion of the ACT when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

The linear relationship between the ACT math versus MAP math scores, stratified by SES, and stratified by Race appeared to have crossing fitted linear regression lines, suggesting a disordinal interaction, and reflecting a significant difference between the slopes of the linear regression lines with respect to SES and Race. Further analysis confirmed the significance of this interaction using ANCOVA.

The interaction with SES was significant and the interaction with Race was also significant. Consequently, the slopes of the regression lines were significantly different. The effect sizes, however, were very low. Statistical evidence was provided to reject one

part of H_05: There is no difference in the measurement of strength of association between the 8th grade MAP math examination compared to the math portion of the ACT when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

The linear relationship between the MAP math versus Algebra I EOC scores, stratified by SES and stratified by Race appeared to have parallel fitted linear regression lines, suggesting no interaction, and reflecting no significant difference between the slopes of the linear regression lines with respect to SES and Race. Further analysis confirmed the lack of significance of this interaction using ANCOVA.

The interactions with SES and Race were not significant and the effect sizes were zero. Consequently, the slopes of the linear regression lines illustrated were not significantly different. There was insufficient statistical evidence to reject the final part of H_05: There is no difference in the measurement of strength of association between the 8th grade MAP math examination compared to Algebra I EOC examination when considering race (minority and not minority) and free/reduced lunch status (non-qualifier and qualifier).

Conclusions

The following conclusions are derived from this study:

1. Results for students that qualified as minority were lower for all the assessments studied.

2. When accounting for socioeconomic status, students consistently performed lower then students that did not qualify for free and reduced lunch status.

3. The lower effect size combined with the significantly lower performance for the identified subgroups of minority and free and reduced lunch indicate that a

larger sample size is necessary to confirm that there is no significant

interaction between Socio Economic Status (SES) and Race

4. Testing format variables have little influence on high stakes assessment scores

    for students taking mathematics assessments.

Recommendations

The following are recommendations for policy change based on the conclusions:

1. Policy makers should be less concerned as to the format of the test that is

    being given, but rather what steps the developer of the test is taking to ensure

    test validity.

2. An opportunity for policy makers is to examine whether policy would be

    beneficial to help ensure students of minority and students that qualify for free

    and reduced lunch have access to supports to help them perform just as well as

    students that do not qualify for these subgroups regardless of test format.

The following are recommendations for educators based on the conclusions:

1. Leaders of school districts in Missouri need to look at developing further

    supports for students that qualify for free and reduced lunch status as well as

    for minority students to help better prepare them for all high-stakes

    assessments.

2. Teachers need to be aware that the test format does not necessarily impact

    student performance, but rather students that qualify as minority or for free

    and reduced lunch tend to perform lower on high stakes assessments.

The following are recommendations for further study:

1. This study could be replicated with a larger sample size to further examine the

    small effect sizes found.

2. This study could be replicated for other subject areas tested within the ACT (English, Reading, Science, and Writing).

3. With students performing differently within the independent variable groups on all three assessments, further study would be warranted into determining the factors that are contributing to these results.

4. Further study as to why the descriptive statistics was significantly lower for the math portion of the ACT and the other two assessments studied is needed.

5. This study could be replicated for other school districts that do not already have a 1:1 initiative set in place for students.

6. With the recent expansion of CBAs being expanded to grades 3 through 8 in the state of Missouri, further study of the impact of CBAs on student performance for those specified grade levels is needed.

7. Further study regarding the impact of different types of CBA formats on student performance is needed.

Limitations of the Study

This study was limited in a number of ways. First, this study included only students that went on to take the ACT test. Students that do not plan on attending a 4-year college often take different assessments. Additionally, not all students that plan to attend a 4-year college take the ACT, but some will take the SAT instead. This study was also limited to high schools in the Kansas City, Missouri area. The largest percentage of students going onto a 4-year college in this area takes the ACT assessment.

Additional limitations were that the EOC exams are used only in Missouri and different types of assessments in different states may have different results. The tests are created independently and different steps may be taken to transition assessments to an

online format. Another study limitation was that students may have been subjected to varying test preparation models that may impact the results outside of the variables being studied. Also, students were sampled from schools that had a 1:1 technology environment within the school. Schools that did not have a 1:1 technology environment were not used within the study. The final limitation of the study was that results of this study described only the population that took these assessments during the 3-year period of high school graduates from 2012-2015 being examined.

Summary

This study sought to add new knowledge to this field of study by determining whether there was a difference in results for students on P&P high stakes assessments and computer-based high stakes assessments when considering race and/or free and reduced lunch status. Through this study, it was established that there is little difference in results for students on P&P high stakes assessments and computer-based high stakes assessments when considering race and/or free and reduced lunch status. Continued research is needed regarding the various formats of CBAs and how the assessments may have an impact on the same established student populations. The results of this study could inform school district leaders and researchers who seek to understand whether a change in assessment format might have an impact on how student results are evaluated. The benefit will come from the increased understanding of how shifts in assessment models could potentially impact students in a high stakes situation.

REFERENCES

ACT. (2014). *Overview*. Retrieved from http://www.act.org/products/k-12-act-test/

Akdemir, O. & Oguz, A. (2008). Computer-based testing: An alternative for the assessment of Turkish undergraduate students. *Computers & Education, 51*(3), 1198-1204. doi:10.1016/j.compedu.2007.11.007

Attewell, P. (2001). Comment: The first and second digital divides. *Sociology of Education*, *74*(3), 252. doi:10.2307/2673277

Barshay, Jill (2016, January). Using computers widens the achievement gap in writing, a federal study finds. *The Hechinger Report.* Retrieved from: http://hechingerreport.org/online-writing-tests-widen-achievement-gap/

Bennett, R. E. (1998, June). *Reinventing assessment. Speculations on the future of large-scale educational testing: A policy information perspective.* Princeton, NJ: Educational Testing Service.

Bimber, B. (1990, May). Karl Marx and the three faces of technological determinism. *Social Studies of Science*, *20*(2), 333-351. doi:10.1177/030631290020002006

Bridgeman, B., Lennon, M. L., & Jackenthal, A. (2003). Effects of screen size, screen resolution, and display rate on computer-based test performance. *Applied Measurement in Education, 16*(3), 191-205. doi:10.1207/S15324818AME1603_2

Brosnan, M. J. (1998). The impact of computer anxiety and self-efficacy upon performance. *Journal of Computer Assisted Learning, 14*(3), 223-234. doi:10.1046/j.1365-2729.1998.143059.x

Carey, L. M. (1994). *Measuring and evaluating school learning* (2nd ed.). Boston, MA: Allyn & Bacon.

Celano, D. & Neuman, S. B. (2010, November). Roadblocks on the information highway. *Educational Leadership, 68(3),* 50-53.

Clariana, R. & Wallace, P. (2002). Paper-based versus computer-based assessment: Key factors associated with the test mode effect. *British Journal of Educational Technology, 33*(5), 593-602. doi:10.1111/1467-8535.00294

Creswell, J. (2003). *Research design: Qualitative, quantitative, and mixed method approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Dean, C. B., Hubbell, E. R., Pitler, H., & Stone, B. (2012). *Classroom instruction that works* (2nd ed.). Denver, CO: Mid-continent Research for Education and Learning.

Duncan, A. (2010). Speech entitled *Education and International Competition: The Win-Win Game. Secretary Duncan's Remarks to the Council on Foreign Relations, New York City*. Retrieved from http://www.ed.gov/news/speeches/education-and-international-competition-win-win-game-secretary-duncans-remarks-council-foreign-relations-new-york-city

Earle, R. S. (Ed.). (2002, January-February). The integration of instructional technology into public education: Promises and challenges. *Educational Technology, 42*(1), 5-13. Retrieved from http://asianvu.com/digital-library/educational_technology/earle.pdf

Editorial Projects in Education Research Center. (2011, September 19). Issues A-Z: No Child Left Behind. *Education Week.* Retrieved from http://www.edweek.org/ew/issues/no-child-left-behind/

Escudier, P. M., Newton, J. T., Cox, J. M., Reynolds, A. P., & Odell, W. E. (2011). University students' attainment and perceptions of computer delivered

assessment; A comparison between computer-based and traditional tests in a "high-stakes" examination. *Journal of Computer Assisted Learning, 27*(5), 440-447. doi:10.1111/j.1365-2729.2011.00409.x

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Fraenkel, J. R. & Wallen, N. E. (2003). *How to design and evaluate research in education* (6th ed.). New York: McGraw Hill.

Garcia, J., Spalding, E., & Powell, R. R. (2001). *Contexts of teaching: Methods for middle and high school instruction.* Upper Saddle River, NJ: Prentice-Hall, Inc.

Great Schools Partnership. (2014). The glossary of education reform. Retrieved from http://edglossary.org/high-stakes-testing/

Hargreaves, M., Shorrocks-Taylor, D., Swinnerton, B., Tait, K., & Threlfall, J. (2004). Computer or paper? That is the question: Does the medium in which assessment questions are presented affect children's performance in mathematics? *Educational Research, 46*(1), 29-42. doi:10.1080/0013188042000178809

Heinich, R., Molenda, M., Russell, J. D., & Smaldino, S. E. (2002). *Instructional media and technologies for learning* (7th ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

Herold, B. (February, 2016). Technology in education: An overview. *Education Week.* Retrieved from http://www.edweek.org/ew/issues/technology-in-education/

Hohlfeld, T. N., Ritzhaupt, A. D., & Barron, A. E. (2013). Are gender differences in perceived and demonstrated technology literacy significant? It depends on the model. *Educational Technology Research and Development, 61*(4), 639-663. doi:10.1007/s11423-013-9304-7

Horton, S. V. & Lovitt, T. C. (1994). A comparison of two methods of administering

 group reading inventories to diverse learners: Computer versus pencil and paper.

 Remedial and Special Education, 15(6), 378-390.

 doi:http://dx.doi.org/10.1177/074193259401500606

Jaschik, S. (2010). *New Evidence of Racial Bias on SAT. Insidehighered.com*. Retrieved

 13 June 2016, from https://www.insidehighered.com/news/2010/06/21/sat

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

 modes of the statewide English test for students with and without disabilities.

 Educational Assessment, 15(2), 107-121. doi:10.1080/10627197.2010.491066

Kim, D. & Huynh, H. (2007). Comparability of computer and paper-and-pencil versions

 of algebra and biology assessments. *Journal of Technology, Learning, and*

 Assessment, 6(4), 31. Retrieved from http://files.eric.ed.gov/fulltext/EJ838614.pdf

Kist, W. (2013, March). New literacies and the common core. *Educational Leadership*,

 70(6), 38-43.

Lee, J., & Reeves, T. (2012). Revisiting the impact of NCLB high-stakes school

 accountability, capacity, resources state NAEP 1990-2009 reading and math

 achievement gaps and trends. *Educational Evaluation and Policy Analysis, 34*(2),

 209-231. doi:10.3102/0162373711431604

Lynch, Mathew. (2015). 4 Ways Digital Tech Has Changed K-12 Learning. *The Journal.*

 Retrieved from https://thejournal.com/articles/2015/05/20/4-ways-digital-tech-

 has-changed-k12-learning.aspx

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

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

 Educational Leadership Journal, 14(4), 117-125.

McDonald, A. S. (2002). The impact of individual differences on the equivalence of

  computer-based and paper-and-pencil educational assessments. *Computers &

  Education*, *39*(3), 299-312. doi:10.1016/S0360-1315(02)00032-5

McDonald, J. H. (2014). *Handbook of biological statistics* (3rd ed.). Baltimore, MD:

  Sparky House Publishing. This web page contains the content of pages 220-228 in

  the printed version. Retrieved from

  http://www.biostathandbook.com/ancova.html

McFadden, A. C., Marsh, G. E., & Price, B. J. (2002). Computer testing in education:

  Emerging trends. *Computers in the Schools, 18*(2), 18.

  doi:10.1300/J025v18n02_04

McKnight, K. (2012). Top 12 ways technology changed learning. Retrieved from

  http://www.teachhub.com/how-technology-changed-learning

Missouri Department of Elementary and Secondary Education, Assessment. (2012).

  Online End-of-Course assessments: Guide to interpreting results 2012-2013.

  Retrieved from

  http://dese.mo.gov/divimprove/assess/documents/asmt-eoc-gir-fall12-spring13.pdf

Missouri Department of Elementary and Secondary Education, Assessment. (2013). MAP

  test information for parents. Retrieved from http://dese.mo.gov/ccr/MAP-info-4-

  parents.html

Missouri Department of Elementary and Secondary Education. (2014). LEA Guide to the

  Missouri Assessment Program. Retrieved from

  http://dese.mo.gov/sites/default/files/asmt-guide-to-missouri-assessment-

  program1415.pdf

Missouri Department of Elementary and Secondary Education, Assessment. (2015). End-of-Course. Retrieved from http://dese.mo.gov/college-career-readiness/assessment/end-course

NAEP writing computer-based assessment: An overview for grades 8 and 12. NCES 2010-470. (2010). Washington, D.C.: National Assessment of Educational Progress (NAEP). Retrieved from https://nces.ed.gov/nationsreportcard/pdf/writing/2010470.pdf

Technology in Schools Task Force (2002). *Technology in Schools. Suggestions, Tools and Guidelines for Assessing Technology in Elementary and Secondary Education.* Washington, D.C.: National Center for Education Statistics (NCES). Retrieved from http://nces.ed.gov/pubs2003/2003313.pdf

Newhouse, C. P. (2011). Using IT to assess IT: Towards greater authenticity in summative performance assessment. *Computers & Education*, *56*(2), 388–402. doi:10.1016/j.compedu.2010.08.023

NOETH, R.J., & VOLKOV, B.B. (2004). *Evaluating the effectiveness of technology in our schools: ACT Policy Report.* Washington, DC: ACT.

Poggio, J., Glasnapp, D. R., Yang, X., & Poggio, A. J. (2005). A comparative evaluation of score results from computerized and paper-and-pencil mathematics testing in a large scale state assessment program. *Journal of Technology, Learning, and Assessment, 3*(6), 31.

ProProfs Quiz Maker. (2014). Defining computer-based assessment. Retrieved from http://www.proprofs.com/c/online-assessment-software/defining-computer-based-assessment/

Public Service Commission of Canada. (2011). Paper-and-pencil instruments. Retrieved

from http://www.psc-cfp.gc.ca/ppc-cpp/acs-cmptnc-evl-cmptnc/pp-instrmnt-pc-

eng.htm

Richardson, W. (2013, March). Students first, not stuff. *Educational Leadership*, *70*(6),

10-14.

Ricketts, C., Wilkes, S. J. (2002). Improving Student performance through computer-

based assessment: Insights from recent research. *Assessment & Evaluation in*

Higher Education, 27(5), 475-479.

Rooks, N. (2012). *Why It's Time to Get Rid of Standardized Tests*. *TIME.com*. Retrieved

13 June 2016, from http://ideas.time.com/2012/10/11/why-its-time-to-get-rid-of-

standardized-tests/

Roschelle, J., Pea, R., Hoadley, C., Gordin, D., & Means, B. (2000, Fall/Winter).

Changing how and what children learn in school with computer-based

technology. *Children and Computer Technology, 10*(2), 76–101. Retrieved from

http://www.futureofchildren.org/information2826/information_show.htm?doc_id

=69809

Rosenfeld, B. (2008). The challenges of teaching with technology: From computer idiocy

to computer competence. *International Journal of Instructional Media*, *35*(2),

157–166.

Ryan, T., & Kellogg, L. (2014, Fall). The final countdown. *EdTech: Focus on K-12,*

12(4), 40-41.

Sadker, M. P., & Sadker, D. M. (2002). A custom textbook prepared for Dr. Leslie

Davison's introduction to middle level and high school education. *Select Material*

from Teachers, Schools, and Society (6th ed.). New York, NY: The McGraw-Hill

Companies, Inc.

Schacter, J. (1999). *The impact of education technology on student achievement: What the most current research has to say*. Santa Monica, CA: Milken Exchange on Education Technology.

Shapley, K. S. (2006, June). *Effects of technology immersion on teaching and learning: Evidence from observations of sixth-grade classrooms*. Austin, TX: Texas Center for Educational Research.

Smith, J. (2014, Spring). Cut to the core. *EdTech: Focus on K-12, 12*(2), 48.

Smith, M. R. & Marx, L. (1994). *Does technology drive history?: The dilemma of technological determinism.* Cambridge, Mass: MIT Press.

Stobart, G. & Eggen, T. (2012). High-stakes testing – value, fairness and consequences. *Assessment in Education: Principles, Policy & Practice*, Vol. 19, Iss. 1, 1-6.

Surry, D.W. & Farquhar, J.D. (May 1997) Diffusion theory and instructional technology. *Journal of Instructional Science and Technology (2)*1. Queensland: University of Southern Queensland. Retrieved from http://www.usq.edu.au/electpub/e-jist/vol2no1/article2.htm

Tschopp, J. K. & Kurdek, L. A. (1981). An assessment of the relation between traditional and paper-and-pencil formal operations tasks. *Journal of Research in Science Teaching, 18*(1), 87-91. doi:10.1002/tea.3660180114

Weglinsky, H. (1998, September). *Does it compute? The relationship between educational technology and student achievement in mathematics*. Princeton, NJ: Educational Testing Service.

Whithaus, C., Harrison, S. B., & Midyette, J. (2008). Keyboarding compared with handwriting on a high-stakes writing assessment: Student choice of composing medium, raters' perceptions, and text quality. *Assessing Writing, 13*(1), 4-25. doi:10.1016/j.asw.2008.03.001

Whiting, G. & Ford, D. (2009). *Cultural Bias in Testing | Education.com*. *Education.com*. Retrieved 13 June 2016, from http://www.education.com/reference/article/cultural-bias-in-testing/

Willhoft, J. (2012, December). Principals leading the way. *Principal Leadership*, *13*(4), 18-21.

Woolfolk, A. (2001). *Educational Psychology* (8th ed.). Boston, MA: Allyn and Bacon.

Young, M. J., Brooks, T., & Olson, J. (2008). Comparability of computer-based and paper-and-pencil testing in K-12 reading assessments. A meta-analysis of testing mode effects. *Educational and Psychological Measurement, 68*(1), 5-24. doi:http://dx.doi.org/10.1177/0013164407305592

Zehr, M. A. (1997). Teaching the teachers. In Technology counts: Schools and reform in the information age. *Education Week, 17,* 24-29.

APPENDIX A

IRB DETERMINATION NOTICE

eCompliance

IRB Determination Notice Project #2005410 Review #214634

Project #2005410
Project Title: A COMPARISON FROM 2008-2015 BETWEEN MISSOURI PUBLIC SCHOOL STUDENT COMPUTER-BASED AND PAPER-AND-PENCIL BASED HIGH STAKES ASSESSMENTS BY RACE AND SOCIO-ECONOMIC STATUS
Principal Investigator: Joshua A Peters, EDD
Primary Contact: Joshua A Peters, EDD

Dear Investigator,

The MU Institutional Review Board reviewed your application and supportive documents. It has been determined that this project does not constitute human subjects research according to the Department of Health and Human Services regulatory definitions. As such, there are no further IRB requirements.

If you have questions, please feel free to contact the IRB office at 882.3181.

Sincerely,

MU Institutional Review Board

APPENDIX B

ACT Results

School District A

Year	# Grads	# Grads at or above National Average	% of Grads at or above National Average	% of Grads Tested	Composite ACT Score
2014	1,391	445	32.00	61.18	21.1
2013	1,311	465	35.50	63.08	21.2
2012	1,366	482	35.30	65.15	21.1
2011	1,335	455	34.10	59.48	21.4
2010	1,389	457	32.90	57.88	21.3
2009	1,270	409	32.20	54.25	21.6
2008	1,225	397	32.40	56.24	21.6

School: A HIGH

Year	# Grads	# Grads at or above National Average	% of Grads at or above National Average	% of Grads Tested	Composite ACT Score
2014	342	86	25.10	54.97%	20.4
2013	331	89	26.90	57.70%	20.3
2012	373	118	31.60	60.05%	20.8
2011	383	115	30.00	52.74%	21.7
2010	389	109	28.00	55.53%	20.7
2009	394	121	30.70	51.52%	21.7
2008	391	104	26.60	50.64%	21.1

School: B HIGH

Year	# Grads	# Grads at or above National Average	% of Grads at or above National Average	% of Grads Tested	Composite ACT Score
2014	359	118	32.90	62.12%	21.2
2013	333	122	36.60	63.66%	21.5
2012	289	104	36.00	62.98%	21.4
2011	294	115	39.10	61.90%	22
2010	300	112	37.30	56.33%	22.4
2009	512	199	38.90	63.09%	21.8
2008	499	210	42.10	65.73%	22.4

School: C HIGH

Year	# Grads	# Grads at or above National Average	% of Grads at or above National Average	% of Grads Tested	Composite ACT Score
2014	357	157	44.00	70.59%	22.2
2013	329	166	50.50	74.16%	22.3
2012	369	179	48.50	81.84%	21.7

Year	# Grads	# Grads at or above National Average	% of Grads at or above National Average	% of Grads Tested	Composite ACT Score
2011	303	148	48.80	78.88%	21.7
2010	301	131	43.50	70.43%	21.7

School: D HIGH

Year	# Grads	# Grads at or above National Average	% of Grads at or above National Average	% of Grads Tested	Composite ACT Score
2014	333	84	25.20	56.46%	20.4
2013	318	88	27.70	56.60%	20.4
2012	335	81	24.20	54.33%	20.3
2011	355	77	21.70	48.17%	20.1
2010	399	105	26.30	51.88%	20.6
2009	364	87	23.90	43.96%	21
2008	335	83	24.80	48.66%	20.7

(DESE, mcds.dese.mo.gov, 2014)

APPENDIX C

EXECUTIVE SUMMARY

COMPARISON BETWEEN A MISSOURI HIGH SCHOOL DISTRICTS' STUDENTS COMPUTER-BASED AND
PAPER-AND-PENCIL BASED HIGH STAKES ASSESSMENTS BY RACE AND SOCIO-ECONOMIC STATUS
by Joshua Peters

Statement of the Problem
This year there will be more state standardized tests given using technology than there will be given via paper-and-pencil. There is a lack of knowledge in whether there is a difference in results for students on P&P high stakes assessments and computer-based high stakes assessments when considering race and/or free and reduced lunch status (Herold, 2016).

Purpose of the Study
The purpose of this study will be to add new knowledge to this field of study by determining whether there is a difference in results for students on P&P high stakes assessments and computer-based high stakes assessments when considering race and/or free and reduced lunch status.

Theoretical Framework
The theoretical framework of this study will be centered on research and theories that pertain to technology and how it affects curriculum and student learning experiences. The theory of technological determinism "claims that technology is the dominant factor in social change" (Bimber, 1990, p. 2).

Conceptual Underpinnings
The conceptual underpinnings supporting the research are technology in schools, CBAs, as well as test anxiety and educational testing.

Research Design
The research design will use a static-group comparison design that utilizes a quantitative, non-experimental correlational methodology.

Research Questions
1. What are the descriptive statistics for student performance on the 8th grade Math MAP test, the Algebra I EOC examination, and the Math portion of the ACT broken out by subgroups?*
2. Is there a difference in the relationship between student performance on the 8th Grade Math MAP examinations when broken out by subgroups?*
3. Is there a difference in the relationship between student performance on the Algebra I EOC examinations when broken out by subgroups?*
4. Is there a difference in the relationship between student performance on the Math portion of the ACT when broken out by subgroups?*
5. What is the difference in the measurement of strength of association between the 8th Grade Math MAP examinations and the Algebra I EOC examination compared to the 8th Grade Math MAP examinations and the math portion of the ACT when broken out by subgroups?*

*Each question will be analyzed by Race: a. Minority and Not Minority, and by Free/Reduced Lunch Status:
a. Qualifiers and Non-qualifiers.

Delimitations
This study focuses on high stakes math assessments in a single school district in Missouri that contained four high schools. All four schools had 1:1 laptops. This study did not look at other curricular areas outside of Math. The study did not explore differences between schools that had 1:1 and schools that did not. Additionally, the study did not look at grades beyond 9-12 and any additional states.

Limitations
Sample size may not accurately reflect all high school students across the state of Missouri. Assessments being compared from paper and pencil to online are from the same student, but not the same test. The sample school district is the fourth largest in the state. The study is limited only to students that took all three assessments within the district and so student mobility impacted sample size.

Significance of the Study
By contributing further to the existing research, the findings of this study may provide educators and policy makers additional information to make decisions that inform future policy discussions regarding the use of CBT for high stakes assessments within the K-12 public education setting.

VITA

Joshua Aaron Peters was born in LaGrange, Illinois to Janet and Leslie Peters on July 15, 1976. Following graduation from Hunterdon Central Regional High School in Flemington, New Jersey in 1994, he received the following degrees: B.A. in Art and Sociology from Luther College (1998); B.S. in Art Education from St. Cloud State University (2004); M.Ed. in Administration and Supervision from University of Phoenix (2009); Ed.D. in Educational Leadership and Policy Analysis from the University of Missouri-Columbia (2016). He is married to the former Aryn Adeline Danner of Kansas City, Missouri. He is father to Cayden Joshua Peters, Emma Adeline Peters, and Grace Catherine Peters of Kansas City, Missouri.