Fall 9-2012

Computer-Based Versus Paper-Pencil Modes of Administration United States Government End of Course Exams: Student Cumulative Grade Point Averages as Predictors of Success

Barbara A. Ryan
Lindenwood University

Computer-Based Versus Paper-Pencil Modes of Administration United States Government End of Course Exams: Student Cumulative Grade Point Averages as Predictors of Success

by

Barbara A. Ryan

A Dissertation submitted to the Education Faculty of Lindenwood University

in partial fulfillment of the requirements for the degree of

Doctor of Education

School of Education

Computer-Based Versus Paper-Pencil Modes of Administration

United States Government End of Course Exams:

Student Cumulative Grade Point Averages as

Predictors of Success

by

Barbara A. Ryan

This dissertation has been approved as partial fulfillment of the requirements for the

degree of Doctor of Education

at Lindenwood University by the School of Education

Dr. Graham Weir, Dissertation Chair	9-14-2012 Date
Dr. Amy Spears, Committee Member	9-14-2012 Date
Dr. Sherrie Wisdom, Committee Member	9-14-2012 Date

Declaration of Originality

I do hereby declare and attest to the fact that this is an original study based solely upon my own scholarly work here at Lindenwood University and that I have not submitted it for any other college or university course or degree here or elsewhere.

Full Legal Name: Barbara Alice Ryan

Signature: Barbara A. Ryan Date: 9/14/12

Acknowledgements

I would like to thank my committee members for all their time, patience, and direction during the dissertation process. Thank you Dr. Weir for agreeing to chair my committee and for providing guidance, feedback, and encouragement. Thank you Dr. Spears for the reassurance I received from you from the first day we met in class as fellow students. Thank you Dr. Wisdom for the numerous hours helping with the statistical compilation and analysis segment of my dissertation. In addition to my committee, I would like to thank Dr. Kania-Gosche for direction throughout the final writing process and Dr. Tyson for his aid as a statistics coach.

Thank you to my friends, colleagues, and students I have taught, who believed in me and galvanized me to complete the dissertation journey. Thanks to my mother who encouraged me to reach for the stars and believed her daughter could do anything to which she set her mind. To my sister Carol, my best friend and a dedicated educator, for allowing me to bend her ear and providing me the voice of reason, while helping me retain my sanity during the dissertation process. Thanks to my awesome children Megan, Jeffrey, and Jennifer, and their spouses, along with my seven grandchildren. I am so proud to be their mom and grandmother and it is because of them my belief in leading by example and in being a life-long learner set me on the dissertation path. To my son-in-law Ryan for his expertise with word processing and for accepting my calls at all hours. Most importantly to my husband Mike whose patience, understanding, and support at home allowed me to pursue this goal. This dissertation would not have been possible were it not for these wonderful people who have entered my life.

Abstract

Beginning with the No Child Left Behind federal legislation, states were required to use data to monitor and improve student achievement. For high schools, the Missouri Department of Elementary and Secondary Education chose End of Course Exams (EOC) to demonstrate levels of student achievement. The policy changed from school choice of paper-pencil or computer-based testing to mandated computer-based testing. This study examined whether this decision best demonstrates the level of student mastery. Using high school EOC test scores for United States Government exams as the independent variable and high school cumulative grade point average (CGPA) as the dependent variable, the study examined the correlation between CGPA and computer-based (CBT) versus paper-pencil (PPT) modes of testing.

Random samples from two comparable school districts were used to provide data. School A tested using computers, while school B used paper-pencil testing. Data presented in this study demonstrate there is little relationship between CGPA and EOC scores depending upon the mode of test administration. For the most part, the null hypotheses were not rejected. Results indicated limited support in some subgroups for the alternative hypotheses that students with a 2.5 or higher cumulative grade point average will score higher on end of course paper-pencil tests, while students with less than a 2.5 cumulative grade point average will score higher on end of course computer-based tests.

Results of this study call into question whether the state and school districts should allow students choice of test mode or perhaps even require students to take the test using

the mode of administration their cumulative grade point average indicates would demonstrate their actual level of achievement. This study also questions whether other high stakes tests such as the ACT, SAT, TOEFL, and LSAT, should determine mode of administration based on students' CGPA.. Finally with the push for data driven classroom curriculum assessment, should the results of this study be applied to the need for differentiation in the classroom with regard to assessments.

Table of Contents

List of Abbreviations

ACT	American College Test
AYP	Annual Yearly Progress
CBT	Computer Based Testing
CGPA	Cumulative Grade Point Average
CLE	Course Level Expectations
EOC	End of Course
FCAT	Florida Comprehensive Assessment Test
GLE	Grade Level Expectations
GMAT	Graduate Management Admission Test
GPA	Grade Point Average
GRE	Graduate Record Examination
HSGPA	High School Cumulative Grade Point Average
LSAT	Law School Admissions Test
MAP	Missouri Assessment Program
MCAT	Medical College Admission Test
MODESE	Missouri Department of Elementary and Secondary Education
NCLB	No Child Left Behind
NICET	National Certification in Engineering Technologies
PPMC	Pearson Product Moment Correlation
PPT	Paper Pencil Testing
SAT	Standardized Admission Test
SPSS	Statistical Package for Social Studies
TOEFL	Test of English as a Foreign Language

List of Tables

List of Figures

Chapter 1: Introduction

Statement of the Problem

In a short period of time the use of computers in education has changed the way students learn and the manner in which student achievement is assessed. Recently high stakes tests, like End-of-Course (EOC) exams for high school students in the state of Missouri, have shifted into computerized methods of test administration. Questions arise as to whether these computerized methods are comparable to the traditional paper-pencil tests, and whether students benefit more from the use of one or the other. As part of the No Child Left Behind (NCLB) Act of 2001, Title 1 of the Elementary and Secondary Education Act of 1965 was amended to read "Title 1 – improving the academic achievement of the disadvantaged" (U.S. Department of Education, 2004, p. 1). Stressing this need for academic achievement, NCLB further stated the importance of all children receiving a high-quality education and being able to demonstrate their proficiency on standardized academic assessments (U.S. Department of Education, 2004). To achieve these goals the sixth of NCLB's 12 proposals called for using state assessments to demonstrate students are meeting "State academic and content standards and increasing achievement overall" (U.S. Department of Education, 2002, Sec. 1001). Even though NCLB mandated annual assessments for all states, the federal government did not mandate the format in which these assessments should be administered. Thus, whether the tests are administered through paper-pencil or computerized methods are left to the discretion of the states.

As these annual assessments are used in "rewarding or sanctioning schools, educators, and students on the basis of test results" (Westchester Institute for Human

Services Research, 2003), they are considered high-stakes tests. For school districts, these test scores are used as a part of school accreditation and for measuring adequate yearly progress (AYP). "School districts must report student EOC (End of Course) exam results and the state uses that information as one goal needed for district accreditation" (Ferguson Florissant School District, 2011, para. 2). "The No Child Left Behind Act (NCLB) of 2001 requires all schools, districts and states to show that students are making Adequate Yearly Progress (AYP)" (Missouri Department of Elementary and Secondary Education, 2010d, p.1). NCLB requires states to establish the following targets by 2014: all students will score at or above the proficient level on state assessments show improvement in attendance and/or graduation rates, and demonstrate a 95% participation rate.

> Missouri's AYP targets were established by the Missouri Department of Elementary and Secondary Education (MODESE) based on a formula from the NCLB Act and an analysis of Missouri Assessment Program (MAP) data, attendance rate data and graduation rate data from prior years. When all targets are met, the requirements of AYP are met. (MODESE, 2010d, p.1).

The researcher felt the method of test assessment was of importance to the outcomes of these EOC, high-stakes tests, and whether or not Missouri would meet NCLB's AYP targets.

Stecher (n.d.) believed positive effects for students taking these high-stakes EOC tests included information about their own skills and knowledge as well as the rewards that come with hard work. Rewards that students enjoy include graduation and higher cumulative grade point averages (CGPAs). Although students are required to take the

EOC tests for graduation, there is no set score necessary to demonstrate passing the exam. “The state requires students to take the EOC exams in order to graduate” (Ferguson Florissant School District, 2011, para. 2). Additionally “DESE strongly recommends that a student’s end-of-course exam score account for a minimum of 10 percent, but no more than 25 percent, of the course grade” (MODESE, 2008b, Question 12).

Missouri educators, in order to comply with NCLB, wrote the Missouri Knowledge and Performance Standards, curriculum frameworks and state mandated assessments. Expectations are that “Missouri students must build a solid foundation of factual knowledge and basic skills in the traditional content areas” (MODESE, 2008c, para. 1). In order to demonstrate this knowledge base, the state of Missouri requires school districts to implement the Show-Me-Standards requiring students to acquire knowledge and skills; gather, analyze, and apply information and ideas; communicate effectively within and beyond the classroom; recognize and solve problems; make decisions and act as responsible members of society (MODESE, 2011a, Overview of Performance Standards, para. 4). “The Show-Me Standards (are) a set of 73 rigorous standards intended to define what students should know and be able to do by the time they graduate from Missouri's public high school” (MODESE, 2008a, The Outstanding Schools Act, para. 2). At each grade level there are distinct standards and sequenced expectations in the forms of Grade Level Expectations (GLEs) for K-8 and Course Level Expectations (CLEs) for 9-12, which are to be incorporated into schools’ curriculums.

> Course-Level Expectations (CLEs) outline the ideas, concepts, and skills that form the foundation for an assessed EOC subject area, regardless of student grade

> level. Because a course such as Algebra I could be delivered in middle school or at any grade level in secondary school, CLEs replace the Grade-Level Expectations (GLEs). Districts can offer courses with different titles that cover the same CLEs. (MODESE, 2010a, p.1).

To aid in writing curriculum frameworks that meet the standards of the Missouri School Improvement Program and the Outstanding Schools Act, MODESE provides school districts assistance with aligning their curriculum to the Show-Me Standards (MODESE, 2011a, "About Office of College and Career Readiness," para. 1). Missouri added "a new assessment system of performance events and multiple choice and short answer questions intended to provide an indication of how well students are meeting the Show-Me Standards and how well they compare academically with other students across the nation" (MODESE, 2008a, The Outstanding Schools Act, para. 4). Missouri's standards, frameworks, and assessments were written to reflect what students should know, and be able to demonstrate mastery of, at each specific grade level.

MAP tests, and later EOC tests, were developed as assessments for Missouri schools, to indicate whether students were mastering state content standards. Initially, Missouri used MAP testing in both elementary and secondary schools. The MAP testing started with the Missouri Outstanding Schools Act of 1993. Missouri chose to use MAP tests to fulfill NCLB's requirement for an annual assessment of students academic progress in their public schools (Barker, n.d., History, para. 1). "The only MAP tests that are actually administered on a regular basis are communication arts, math and science" (Barker, n.d.,Considerations, para. 1). After starting MAP testing in 2006 for both

elementary and secondary schools, MAP testing continued in elementary schools, while in 2009 secondary schools transitioned to EOC testing (McGraw Hill, 2009).

"EOC assessments are criterion-referenced tests that are delivered to middle and high school students when the CLEs for a particular course have been covered. English II, Algebra I, Biology, and Government are required EOC assessments for all students to satisfy the requirements of NCLB and the Missouri State Board of Education" (MODESE, 2011b, About the Assessments, para. 1).

The reasons the state decided to move from MAP to EOC testing at the secondary level were based on several factors. First, "in the past, there were no consequences for students if they scored poorly on a MAP test" (MODESE, 2009a, "State Officials Pleased", para. 9). Secondly, allowing school districts to use the EOC test as a semester final or as a course test grade was intended to increase student motivation to attain better grades on the test (MODESE, "State Officials Pleased," 2009a, para 9). In the past, with MAP testing, students saw no reason to try their best on the tests, as the tests had no impact on their grades (Tran, 2009). Students' course grades, CPGAs, and applications to postsecondary institutions would be effected with the advent of EOC testing, thereby motivating students to perform well on their EOCs.

Since all states are required, according to NCLB, to measure a child's progress in reading/language arts and math, in each of Grades 3- 8, and once during Grades 10-12, and in science, once during Grades 3-5, Grades 6-9, and Grades 10-12; previously used Missouri tests were not appearing to demonstrate true evidence of mastery (U.S. Department of Education, 2003). Therefore, the state responded to this federal testing mandate by instituting EOC exams, required annually by secondary schools. States may

require students to undergo assessments in other subject areas, as Missouri did in Social Studies, with the U.S. Government EOC. The Missouri State Board of Education listed the following purposes for EOC assessments: "Measuring and reflecting student's mastery toward post-secondary readiness, identifying students' strengths and weaknesses, communicating expectations for all students, serving as the basis for state and national accountability plans, evaluating programs" (MODESE, 2010a, p. 2).

During the first two school years of testing, 2008-2009 and 2009-2010, EOC tests were available to school districts in either paper-pencil or computer-based format. For the 2010-2011 school year, MODESE required districts to administer the EOC exams via computer.

Mandated CBT created several challenges for school districts, including logistical issues. One concern to high schools was that, during the period designated for testing, almost all computers in a school could be dedicated to EOC testing. As a result, during EOC testing very few school computers would be available for classroom lessons. The state testing window for the fall of 2009 was October 13, 2009, through January 29, 2010, and for the spring of 2010 was March 1, 2010, through May 28, 2010 (MODESE, End-of-Course Assessment, Online Test Examiner's Manual, 2009b, p. 1). "Districts choose one week inside the EOC administration window," to administer the test (MODESE, End-of-Course Assessment Test Coordinator's Manual Training, 2009c., p, 41). School districts have different levels of technology; therefore, some districts need to use the entire testing window because they do not have enough computers for all students to take the test at the same time. As a result of the testing window occurring at the end of the semester, EOC testing overlapped administration of finals for both school districts

in this study, as was the case for some high schools in Missouri. This overlap resulted in computers designated for EOC testing not being available for usage during finals, whether for the finals test itself or for preparation for other types of final projects such as research papers, mock trails, or graphic novels.

High-stakes Testing – Test Modes

Virtually all institutions of learning have some type of testing, including high stakes tests, as a part of their assessment process. High stakes tests are "a single assessment that is given with the knowledge that important decisions or consequences are riding on the result" (Morin, n.d., p.1). "A decision that is primarily made based on the results of a standardized test is called high stakes testing" (Meador, n.d., p.1). "In education, these decisions often relate to federal and local funding, placement and graduation decisions or ongoing tenure for teachers" (Morin, n.d., p. 1). Throughout the country, there is a wide variation in testing modes including computer based testing (CBT), paper pencil testing (PPT), or student choice of testing mode. A large body of research exists regarding the benefits and disadvantages of CBT versus PPT administration. Benefits and disadvantages of modes of testing will be outlined on in Chapter 1 and discussed in greater depth in Chapter 2 of this dissertation.

Some high stakes tests, such as the Medical College Admission Test (MCAT), have changed from paper-pencil to computer-based administration, based on the perceived benefits. Kaplan Test Prep (n.d.) cited MCAT's "desire to offer students greater flexibility: more test sites, more testing times, greater security, a more controlled testing environment, a shorter test day, and faster score results" as reasons for their switch to CBT. However, assessments such as the American College Test (ACT) and Scholastic

Aptitude Test (SAT) continue to use paper-pencil application. Other major examinations including the Graduate Records Exam (GRE), the Praxis Series: Teacher Licensure and Certification Exams, and the Test of English as a Foreign Language (TOEFL), are offering test takers their choice of electronic or paper formats. Some high stakes testing organizations have found it necessary to stay with PPT due to lack of computers, while others have converted to CBT for ease of grading and reporting (Slocum, 2009). Neither reason would be recognized as ensuring best educational practice. Offering a choice of testing mode, on the other hand, would allow test takers to choose the test format they feel will maximize their comfort level and, thus, best demonstrate their subject knowledge (Slocum, 2009).

Compatibility of student learning styles with modes of testing is another issue that has not often been discussed in the educational community. This involves whether students should have choices between CBT and PPT, particularly if CGPA might indicate that one format could more accurately measure their level of mastery. Advocates of best practices in education stress the need for differentiation in the classroom. Kingore (2005) believed in the importance of differentiated student instruction. "Respect for individual differences among and between learners" is Kaplan's "definition of differentiation" (Northwestern University, 2010, para. 9). Questions arise as to whether standardized testing models clash with the need for differentiation to meet students' needs. If, in Missouri, every student is expected to pass the same test, in the same year or same grade, with the same mode of administration, without regard to individual student needs; perhaps there is a need to study the importance of differentiation in testing.

Background of Study

While completing a self directed final in a high school sociology class taught in School District A, one of the two school districts in this study, taught by the author of this study, a pair of students chose to investigate computer-based versus paper-pencil testing. They randomly selected a class of students and had them take two tests on the same topic, the first using paper-pencil and the second using computers. On both the CBT and PPT students were asked to give their CGPA. Results from this project led them to conclude that CGPA was in some way related to how well students had scored on their tests, whether given by computer-based or paper-pencil administration. Although this study was a fledgling attempt at using the scientific method to prove or disprove a theory, the students documented a relationship between CGPA and test format. Their results seemed to indicate that the higher a student's CGPA, the better he or she did with PPT. The lower the student's CGPA, the higher his or her scores seemed to be with CBT. Should this hypothesis hold true upon evaluation of the state-level EOC exams, then perhaps the state mandated CBT only policy might need reevaluation.

One of the current best practices in teaching is the use of differentiation in the classroom. If teachers are to provide students a variety of ways to learn in order to improve student achievement, students should also be provided differentiation in the area of testing. In this era of high stakes testing, in which school districts are required to make AYP, students should take tests in the manner most conducive to achieving their best scores, thereby accurately measuring their mastery of the curriculum. Aspiring to keep or gain accreditation, school districts would benefit from being allowed to implement testing systems which give students the best chance to demonstrate academic progress.

"Districts that do not earn accreditation may receive additional funds and support from the state to implement improvement plans. Additionally, low-performing schools may be eligible to receive federal and/or state money" for such things as tutoring (Great Schools, n.d.a, para 2). High stakes testing consequences for schools or districts could include "public reporting—with its attendant possibility for public praise or censure—to financial rewards for good performance, to a complete state takeover for persistent bad performance" (Barth & Mitchell, 2006, para. 4). In many cases, federal, state, local, and school resources were allocated to low performing schools based on the results of EOC testing as a reflection of student achievement. Districts which fail to make AYP face sanctions that may not be deserved or necessary if student performance levels reflect a mismatch between test taker and mode of testing, rather than lack of subject matter mastery. If students are not tested in a manner conducive to doing their best on EOCs, this might lead to more districts not meeting AYP.

Purpose of the Study

Two purposes were pivotal to this study. One was to determine if CGPAs were a predictor of student success on computer-based and paper-pencil exams, the second purpose was to provide information to the Department of Elementary and Secondary Education, as well as school districts, as to the benefits of one type of test administration over the other, or the benefits of providing both options. Since the present testing mandate for computer-based administration may not provide accurate evaluation of student achievement, this investigation was intended to provide important feedback.

The focus of this study was high school EOC exams for United States Government courses in two suburban school districts in St. Louis County, Missouri. Data used in this

study were collected for the 2009-2010 school year, during which time MODESE allowed school districts to choose between CBT and PPT. The data from this school year was vitally important in order to determine if there was a difference between performance measured by CBT and PPT, as related to CGPAs, since as of the 2010-2011 school year DESE mandated all districts would go to CBT for EOC exams. Thus, after the 2009-2010 school year, there would be no means of collecting data for comparison. Of the two districts used for this study, School District A chose CBT, while School District B chose the PPT test mode for 2009-2010.

There were several reasons the focus of this study was solely on U.S. Government course content. First, the investigator has a background in the field of social studies; therefore, these exams were familiar to the investigator. Second, the investigator had access to two school districts that administered the U.S. Government test to juniors. Not all school districts administer this test in the junior year. Third, using just two school districts and one EOC test, United States Government, provided a more than adequate sample size for testing the validity of the study's hypotheses. Fourth, in addition to limiting the number of study variables, focusing on the 2009-2010 End of Course United States Government exams, required for all students in United States Government courses in the State of Missouri, seemed logical because this was the last school year MODESE allowed districts to choose between CBT and PPT for all EOC testing.

Numerous computer-based versus paper-pencil testing comparability studies are available; but very few, if any, have evaluated the connection between CGPA and type of testing administration. The project completed by the two previously mentioned Sociology students indicated that students with higher CGPAs scored higher taking

paper-pencil tests, while students with lower CGPAs scored higher on computer-based tests. As a result of this apparent correlation, a second purpose of the study was to evaluate whether MODESE's mandate that, beginning in 2010-2011, all end of course testing was to be administered via computers was fair to students and school districts. Data gained from the study might provide valuable information for MODESE to suggest whether the mandate for computer-based administration should continue or be revised to allow school districts, individual schools, teachers, or even students a choice in method of administration to evaluate students' performance on EOC exams. Since, in addition to measuring individuals' mastery and achievement, assessments provide diagnostic information as to what areas have been taught most effectively and which areas need further attention, another purpose of the project was to determine what approach to testing, CBT versus PPT administration, would provide the most accurate information for improving instruction.

Rationale for the Study

The rationale for this study was to ensure valid assessment of student learning. For EOC testing to fulfill its intended purposes—holding schools accountable to educate all students, as provided by NCLB, and accurately evaluating curriculum and instruction—students should take the tests in a manner that provides evidence of true mastery levels, without distortion due to method of test administration. Therefore, the hypothesis of this study was that students should be evaluated in the manner that best matched their learning style.

In order to help districts assess their progress towards meeting AYP, they need to have an accurate measurement of student proficiency. The Missouri State Board of

Education provided EOC tests for this purpose. In order to achieve valid results, testing conditions must not favor or handicap students due to testing format. It was therefore important to determine whether CGPAs might suggest that different students should have access to different modes of testing.

Questions Addressed in the Study

The following questions were addressed in the study:

1. What is the relationship between students' cumulative grade point averages and their exam scores for United States Government End-of -Course exams, through paper-pencil or computer-based administration?
2. What is the relationship been students' cumulative grade point averages and their exam scores for United States Government End-of-Course exams, through paper-pencil or computer-based administration, disaggregated by No Child Left Behind subgroups?
3. Should the Missouri Department of Elementary and Secondary Education continue with their mandate for computer-based administration of End-of-Course exams, or would choice of mode of administration be in the best interest of students; better demonstrating student mastery of the curriculum as determined by their performance on these exams, and therefore benefit their district and the state?

Independent Variable

High school EOC test scores for United States Government exams were used as the independent variable in this study.

Dependent Variable

The dependent variable used in this study was student high school CGPA, as measured at the end of the 2009-2010 school year. As this is a common measurement of student performance in the United States, and had shown possible correlation with test administration preference, it was used as the dependent variable.

Hypotheses

Null Hypothesis 1: For students taking the United States Government EOC examination by paper-pencil mode of delivery, compared to students taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam.

Null Hypothesis 2: For students taking the United States Government EOC examination by paper-pencil mode of delivery, there is no relationship between student cumulative GPA and United States Government EOC exam score.

Null Hypothesis 3: For students taking the United States Government EOC examination by computer-based mode of delivery, there is no relationship between student cumulative GPA and United States Government EOC exam score.

Null Hypothesis 4: For students taking the United States Government EOC examination by paper-pencil mode of delivery, students with a cumulative GPA of 2.5 or higher will not score higher than students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

Null Hypothesis 5: For students taking the United States Government EOC examination by computer-based mode of delivery, students with a cumulative GPA below 2.5 will not score higher than students with a cumulative GPA of 2.5 or above, on

the United States Government EOC exam.

Alternative Hypothesis 1: For students taking the United States Government EOC examination by paper-pencil mode of delivery, compared to students taking the examination in computer-based mode of delivery, there is a difference in average students' scores on the United States Government EOC exam.

Alternative Hypothesis 2: For students taking the United States Government EOC examination by paper-pencil mode of delivery, there is a relationship between student cumulative GPA and United States Government EOC exam score.

Alternative Hypothesis 3: For students taking the United States Government EOC examination by computer-based mode of delivery, there is a relationship between student cumulative GPA and United States Government EOC exam score.

Alternative Hypothesis 4: For students taking the United States Government EOC examination by paper-pencil mode of delivery, students with a cumulative GPA of 2.5 or higher will score higher than students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

Alternative Hypothesis 5: For students taking the United States Government EOC examination by computer-based mode of delivery, students with a cumulative GPA below 2.5 will score higher than students with a cumulative GPA of 2.5 or above, on the United States Government EOC exam.

Limitations of Study

Validity of Cumulative Grade Point Average

Although CGPA was used as a variable, this study did not take into account how the students earned their CGPAs. Because CGPA is a mathematical computation totaling

the grades earned in each course and dividing by the total number of courses taken, it does not weigh the wide variation in students' course selection represented by the CGPAs used in this study. In other words, one student may take a set of courses considered more advanced or rigorous than another student, but have the same CGPA.

A further limitation of the study was that both school districts A and B allow students some elective choice in course selection. While in high school, students in Missouri have a set number of graduation credits they need to attain, in specific courses, in order to graduate. Among the 24 credits required for graduation in Missouri are 4 credits of Communication Arts; 3 credits each of Social Studies, Mathematics, and Science; 1 credit each of Practical Arts, Fine Arts, and Physical Education; .5 credits each of Personal Finance and Health Education; and 7 elective credits (MODESE, 2007, p. 5). "Elective units are additional offerings which are needed to complete" the total number of credits required for graduation (Monroe County Schools, 2001, 1.0, c, 1, a, 3). A students' choice of these seven elective credits can impact their CGPA.

In addition, within their classes, students in this study may be in different levels of programming including foundations (lowest level courses), regular, and honors classes. While expectations are sometimes different for students enrolled in the same course; but at different levels, all grades are calculated the same. For instance, in the area of U.S. Government, students may be in a regular U.S. Government class, a co-taught U.S. Government class, or Advanced Placement Government (honors).

> Co-teaching is a unique blend of direct and indirect services in which a general educator and a special educator jointly instruct pupils in a single classroom. Co-teaching occurs when two or more professional jointly deliver substantive

> instruction to a diverse, or blended, group of students in a single physical space. (Jackson School District, n.d., p. 1)

As a result, all students in the co-taught class receive the advantage of having extra adults in the classroom. Foundation courses and co-taught classes give remedial help by teaching skills and strategies for raising a student's achievement level. These remedial classes "teach students what they should already have learned" (Education.com, n.d., para. 1).

Yet another limitation of the study was individual teachers' differing grading practices, which result in what Montgomery (2009) termed the "unscientific nature of the grading process" (para. 8). Teachers teaching the same course often use different grading scales, lack objectivity or reliability in essay grading, and differ in weight for certain assignments and total number of points possible in a course. Even with these limitations the investigator felt CGPA was a valid choice for the dependent variable. "Although they evaluate the entire application to make acceptance decisions, college admissions officers usually weigh a student's grade point average (GPA) and SAT scores most heavily" (Clipper, 2010, para. 1). Troseth (2008) agreed, saying "the best indicators for success are the student's grade point average (GPA) and college entrance exams" (p. 1). Relied upon as a measure of student performance in the United States, CGPA provided the most accessible and appropriate dependent variable for the purposes of this study. The study was designed to investigate the possible correlation between CGPAs and preference for testing format, as suggested by the previously mentioned sociology project.

Sample Size

Some school districts require U.S. Government in the ninth grade, while other

districts teach U.S. History in ninth grade and U.S. Government in 10th or 11th grades. An additional limitation of this study was that the sample size used in the study represented only two suburban school districts, both of which used eleventh graders for U.S. Government EOC testing. Student maturation level and the fact that U.S. History, when taken prior to U.S. Government, provides students a strong backbone and supporting knowledge for United States Government should be considered for future studies, if ninth grade U.S. Government EOC scores were to be compared to 11th grade scores.

Both districts in the study were St. Louis County suburban districts and were very similar demographically; thus various groups represented in these districts may differ from those in any future studies, particularly as results are limited because there was limited representation of urban or rural student populations. For this study the students were essentially the same age, as they were 11th graders taking the U.S. Government EOC.

Level of Analysis

This investigation analyzed only data at the aggregate test level to see if CBT and PPT administered test results are interchangeable. One limitation of the study was it did not examine item-level analyses to determine performance under the two modes of administration.

Participant Factors

Although this study examined participant factors such as demographics, CBT, and PPT, other factors were not included. The investigation did not attempt to measure the possible impact of test anxiety, anxiety related to the use of technology while using

computers in the test taking situation, familiarity with computers, typing skills, examinee motivation, or test takers' cognitive aptitudes.

Computer Characteristics

Many different types of computers exist today, with various processing powers and capabilities, as well as various sizes, shapes, and available programs. This study did not attempt to account for differences among computer models utilized, and thus varying characteristics such as screen size, font, laptop or desktop, resolution, and speed as suggested by Wang and Shin (2009), could make a difference in CBT. Studies which have assessed correlations between type of test administration and computer characteristics will be discussed in Chapter 2.

Definition of Terms

Adequate Yearly Progress (AYP). "The No Child Left Behind Act (NCLB, 2002) builds upon the accountability provisions in the Improving America's Schools Act of 1994 (IASA, 1994), which required each state to establish challenging content and performance standards and to implement assessments that measure students' performance against those standards" (Goertz, n.d., p. 1). According to Elmore and Rothman (1999) as cited in Goertz, n.d.), the IASA defined adequate yearly progress (AYP): "In a manner that 1) results in continuous and substantial yearly improvement of each school and local education agency sufficient to achieve the goal of all children … meeting the state's proficient and advanced levels of achievement; [and] 2) is sufficiently rigorous to achieve the goal within an appropriate timeframe" (cited in Goertz, n.d., p. 1). Goertz (n.d.) explained that the NCLB legislation made several critical changes to the IASA definition for

AYP and required each state to create its own definition of AYP within the parameters set by Title I. NCLB stated that each state was required to define AYP in a manner as follows: "(i) Applies the same high standards of academic achievement to all public elementary school and secondary school students in the State; (ii) is statistically valid and reliable; (iii) results in continuous and substantial academic improvement for all students; (iv) measures the progress of public elementary schools, secondary schools and local educational agencies and the State based primarily on the academic assessments ... ; (v) includes separate measurable annual objectives for continuous and substantial improvement for each of the following: (I) The achievement of all public elementary school and secondary school students. (II) The achievement of—(aa) economically disadvantaged students; (bb) students from major racial and ethnic groups; (cc) students with disabilities; and (dd) students with limited English proficiency" (U.S. Department of Education, 2002, Part A, Subpart 1, Sec. 1111, 2[c]).

Computer-based testing (CBT). "A Computer-Based Assessment (CBA), also known as Computer-Based Testing (CBT), e-assessment, computerized testing and computer administered testing, is a method of administering tests in which the responses are electronically recorded, assessed, or both" (Computer-Based, 2010, para. 1).

Co-Taught.

Co-teaching, or having two teachers in the classroom has become a popular teaching structure to provide an inclusive setting for special education students while insuring that they are in the least restrictive environment as recommended

by their IEP team. In the co-teaching classroom there is typically a general education teacher and a special education teacher in the classroom. While co-teaching, both teachers are intended to share the teaching responsibility in a co-teaching classroom, with the special education teacher, providing specialized differentiated lessons for students with special needs (Williams, 2009, Co-teaching in the Classroom, para. 1-2).

Course Level Expectations (CLEs). "The Social Studies Grade- and Course-Level Expectations outline related ideas, concepts, skills and procedures that form the foundation for understanding and learning social studies (MODESE, n.d., p. 1).

Cumulative Grade Point Average (CGPA). "A Cumulative Grade Point Average is the mean GPA from all academic terms within a given academic year, whereas the GPA may only refer to one term" ("Grade," 2010, para. 1).

Dependent Variable. The dependent variable is the variable that cannot be controlled or manipulated (Bluman, 2008).

Differentiation.

Differentiated instruction applies an approach to teaching and learning that gives students multiple options for taking in information and making sense of ideas. Differentiated instruction is a teaching theory based on the premise that instructional approaches should vary and be adapted in relation to individual and diverse students in classrooms (Hall, Strangman, & Meyer, 2003, p. 2).

End of Course Testing (EOC). Sometimes referred to as end of course exams, the Missouri Department of Elementary and Secondary Education states these exams will "provide a valid and reliable method for assessing students' knowledge of

Missouri's Course-Level Expectations (CLEs)", while allowing "classroom teachers to incorporate statewide assessment results into students' course grades" (MODESE, 2008b, Question 1, para. 2).

Grade Level Expectations (GLE). "The Social Studies Grade- and Course-Level Expectations outline related ideas, concepts, skills and procedures that form the foundation for understanding and learning social studies" (MODESE, n.d., p. 1).

Grade Point Average (GPA). "Grade point average (GPA) is calculated by dividing the total amount of grade points earned by the total amount of credit hours attempted" ("How," n.d., para.1). In the high schools used in this study grade point averages range from 0.0 to a 4.0, where A=4.0, B=3.0, C=2.0, D=1.0, F=0.

High Stakes Testing.

Tests are considered high-stakes tests when decisions made based on these test scores, have important consequences for the test taker. Some types of decisions made based on high-stakes testing are: high school graduation, promotion to the next grade, access to resources and special opportunities, and summative measures of teacher quality (Pearlman, 2001, p. 1).

Examples of current high stakes tests are: "College and graduate school Student admissions tests, licensing and certification tests for jobs and professions and increasingly, student tests for K-12 students" (Pearlman, 2001, p. 1).

Independent Variable. The variable in a study that can be controlled or manipulated.

Measures of Academic Progress or Missouri Assessment Program (MAP) Testing. Both the Missouri Assessment Program and the Measures of Academic Progress

contain a series of assessments that students take to demonstrate their progress in the areas of math, reading, using language and science. The Missouri Assessment Program is a series of standardized tests produced and graded by the Missouri Department of Elementary and Secondary Education for students in grades three through eight who attend Missouri public schools (Barker, n.d., para. 1 , Identification).

Mode Effect. The differences found between PPT and CBT test modes. (McClarty & Davis, 2006).

No Child Left Behind (NCLB). According to federal legislation passed in 2001, the purpose of the No Child Left Behind Act was "to close the achievement gap with accountability, flexibility, and choice, so that no child is left behind "(Public Law 107-110, 2002,).

Paper-pencil testing. "Paper and Pencil testing is available for traditional classroom situations, where computer access is limited or where a controlled testing environment is required" ("Insight Assessment," n.d., para. 1).

Summary

Based on the results of a student high school Sociology course final investigation into the relationship between CGPA and test scores, the investigator took these initial results and further tested them, to determine whether permitting students to take the mode of test, which their CGPA indicated would best suit them, would produce more valid test scores. The purpose of this study was to examine the possible relationship between students' CGPAs and their performance on United States Government EOC exams, as influenced by mode of test administration. If test format were shown to affect resulting

exam scores in a manner that correlated with test takers' CGPAs, then perhaps students should have a choice of formats for high stakes tests. Educators tout the importance of differentiation in the classroom. Tomlinson (as cited in Dahlman, Hoffman, and Brauhn, 2011) indicated that differentiated instruction "has proven to be successful in the general education context where studies have found that students exposed to Differentiated Instruction strategies consistently outperform other students" (Abstract, para. 1). Providing students a choice of CBT or PPT is one method of allowing for differentiation. This study analyzed overall test performance, as well as disaggregated test performance of students with CGPAs of 2.5 or higher or students with lower than 2.5 CGPAs.

Little to no research exists connecting CGPA to test performance and how the latter is influenced by method of testing. This study fills a void in this research and gives some insight into CBT and PPT as they may influence results on the new high-stakes tests states are using as part of their accountability under NCLB. Although little to no research exists correlating CGPA with test modes, the literature review in Chapter 2 will provide enlightenment regarding factors that influence both CGPA and methods for administering high stakes tests.

Chapter 2: Review of the Literature

Researching the relationship between CGPA and modes of test administration was the purpose of this study, in order to see if differentiation in testing modes and allowing students' choice of test mode on high stakes tests would present a more accurate reading of their curriculum mastery. Students are required to take many high-stakes tests that will determine the direction of their future, including such items as: grade promotion, graduation, if and what college they will attend, what career they might pursue, and what level their future earning power will encompass. This chapter focused on CGPA and its relationship to CBT and PPT.

A literature review based on research using CGPA as a predictor of success on EOC tests, or any high stakes tests, administered through the use of computer technology versus paper-pencil, yielded very few references. In addition this review examined the following topics with regard to their importance to test mode administration: NCLB; the importance of high-stakes testing; the purposes of EOC testing; comparisons of CBT and PPT; differentiation in the classroom; brain functions reading on line versus on paper; score comparability of assessments on-line versus on paper; and the relationship of CGPA to computer-based or paper-pencil test administration.

GPA as a Predictor of Success

In a paper presented at the Annual Conference of the Mid-South Educational Research Association, Nejad (1995) reported on the effects of college GPA on learning electronics through computer usage and traditional methods. "Results indicate that age nor GPA is a factor in learning electronics via computer simulation-based or traditional breadboard instruction" (Nejad, 1995, p.16). Analyzing undergraduate GPA (UGPA) as

a predictor of success in graduate school, using a students' graduate GPA (GGPA), Carpenter's (2005) results demonstrated that UGPA is not a valid predictor of student success in college judged by GGPA (p. 14). When examining predictors of success for college students taking teacher certification exams, researchers established that "High School GPA was not a strong predictor of future success on a teacher certification exam" ("Teacher Certification Exams," 2005, Findings, para. 8). Although many studies have found no relevance for GPA as a predictor of success, a similar number of studies existed demonstrating just the opposite.

Brown (n.d.) of the National Collegiate Athletic Association (NCAA) stated "high school GPA may be a much stronger predictor of first-year performance in college than standardized tests" (para. 2). "You start to get this indicator that high school GPA is more closely associated with CGPA because you are comparing apples to apples," according to Kevin Schriver (Brown, n.d., para. 7). Geiser and Santelices (2007) discovered "high school grade point average (HSGPA) is consistently the strongest predictor of four-year college outcomes for all academic disciplines, campuses and freshman cohorts in the UC sample" (p. 1). "A statistically significant correlation between cumulative GPA and retention," emerged in a study by DeBerard (2004, p. 6). Yet another study found "college admissions officers usually weigh a student's grade point average (GPA) and SAT scores most heavily. A stronger GPA, with an overall solid application, usually increases a student's chance for admittance" (Clipper, 2010, para. 1). Regarding GPA as a measurement of success in an Introduction to Computers course; Baxter, Hungerford, and Helms discovered "GPA to be a better predictor of the final course grade" (2010). Desmarais, Woble-Valenski, and Oestmann (2011, p. 36)

found "PTA GPA (physical therapist assistant coursework GPA) was the best predictor for success on the NPTE-PTA (national licensure examination for physical therapist assistants) based on higher examination scores" (p. 36). As a predictor of success on the Anesthesiology In-Training Examination (ITE), Warrick and Crumrine (1986) pronounced that "Medical school GPA appears to be an indicator of success on the ITE" (p. 594). "Generally speaking, graduate schools find GPA to be the most reliable predictor of success in graduate school" ("I Feel Like," n.d., "Reliability of GPA," para. 1). GPA is widely accepted as an indicator of academic student achievement.

Besides considering the pros and cons for use of GPA as a predictor, there arises the issue that GPA is not standardized across the United States.

> High schools don't use the same GPA scale – and even when they do, many used weighted systems (perhaps giving extra 'points' to grades from honors, accelerated, International Baccalaureate, or Advanced Placement classes), and employ varying methods of calculating a cumulative GPA ("College Admission Requirements and Your GPA," n.d., "The GPA and College," para. 1)

Even though there are a variety of ways to calculate GPA the literature demonstrated GPA is still one of the best methods for predicting student success.

CGPA is one measure used to decide which students would be considered high-achieving. Clariana and Wallace (2002) in their investigation of factors influencing success on CBT or PPT discovered that "higher-attaining students benefited most from computer-based assessment relative to higher-attaining students under paper-based testing" (p. 593). Their reasoning for these benefits is that "higher attaining students likely accommodated more quickly and so benefited from computer-based assessment"

(Clariana & Wallace, 2002, p. 601). According to Clariana and Wallace (2002) they found a 2001 study by Watson that looked at computer-aided learning (CAL) and perceived that "students with higher academic attainment . . . benefited most from CAL" (p. 594). Thus it is possible to surmise that possibly CGPA, as a measure of high attainment, could be used as a predictor of success with regard to particular test taking modes.

No Child Left Behind

NCLB Act of 2001 became law in 2002 and since that time has impacted schools across the United States. The purpose of the law was to improve academic achievement through what students are taught, what tests they take, what training teachers receive and in what way money is spent on education (Great Schools, n.d.b) "Although the Act mandates annual testing for all states by 2005-2006, it does not provide federal standards for testing practices" (Wenning, Herdman, Smith, McMahon, & Washington, 2003, p. 2). NCLB provided states some testing flexibility in several areas including the following: whether to use norm-referenced or criterion-referenced tests, what subjects will be tested and in which years, and the definition of proficient (Wenning et al., 2003, p. 2). "Norm-referenced tests assess a student's broad knowledge . . . Criterion-referenced tests measure specific skills" (Wenning et al., 2003, p.2). As an example of NCLB impact, in the states push to meet the mandatory annual testing requirements of NCLB, the choice of CBT or PPT is part of the testing practices left to the discretion of each state. In Missouri the Department of Education chose to move to CBT. The idea of this study started with the author's experience in a school district changing testing from PPT to CBT. Because of NCLB, the state uses EOC scores as evidence of academic

improvement. These test scores became a part of state targets to demonstrate AYP under the NCLB guidelines.

Importance of High-Stakes Testing

. High-stakes tests are standardized tests having "consequences attached to the results," which can include issues for students such as grade promotion, graduation, and admittance to college (Barth & Mitchell, 2006, "What Makes a Standardized," para. 1). Sireci (2009), director of the Center for Educational Assessment in the School of Education at the University of Massachusetts, Amherst, discussed the reasons for standardized testing which include the following: accountability by teachers, schools, and districts; requirement for graduation; qualifying for scholarships; participating in athletics; assigning to grade levels; improving student learning; and aligning teaching with state curriculum frameworks. Currently, Missouri's EOC United States Government test, the test used in this study, meets the definition of both a high-stakes test and a standardized test. As of the 2010-2011 school year all EOC tests were required to be CBT, meeting the definition of standardized testing,

> Standardized tests are large-scale tests that are administered to students and scored in the same manner. Students take the same test in the same conditions and, if possible, at the same time so that results can be attributed to student performance and not to differences in the test or the way it is given. Because of this, the results of standardized tests can be compared across schools and districts. (Barth & Mitchell, 2006, "What are Standardized Tests," para. 1)

In the investigator's study, two school districts gave an EOC high stakes test using different modes of administration; therefore this study looked at the impact of

differences in testing mode. Were Missouri to offer a choice of CBT or PPT they would no longer meet the definition for standardized tests, as read previously in Barth and Mitchell's (2006) definition of standardized tests, which includes using the same mode of administration and were required for graduation, meeting the requirement for a high-stakes test.

Why are high stakes tests such an important part of the current educational climate? Reasons given by Nichols and Berliner (2008) include the following:

> High-stakes testing is the practice of attaching important consequences to standardized test scores and it is the engine that drives the No Child Left Behind (NCLB) Act. The rational for high-stakes testing is that the promise of rewards and the threat of punishments will cause teachers to be more motivated, and schools to run more smoothly — all of which will result in greater academic achievement for all students, but especially those from poverty and minority backgrounds. (p.672)

Sloane and Kelly (2003) discussed the RAND publication, Making Sense of Test-based Accountability in Education edited by. Hamilton, Stecher and Klein, where in Chapter 4: Consequences of Large-scale, High-stakes Testing on School and Classroom Practice, were found Stecher's thoughts on the effects of high-stakes testing for students. On the positive side, this type of testing gives students information about their own knowledge and skills, indicates their weak areas of study, encourages them to work harder in school, shows them that putting effort into their studies will reap rewards, and teaches them about competitiveness, which exists in the real world. On the other hand, students

become frustrated and discouraged and begin to loose interest in grades and school assessments.

Sentance (2000), Education Policy Advisor to the Governor of Massachusetts and a member of the Education Commission of the States, answered the question of why there is a need for high-stakes testing by discussing the old educational system under which students weren't learning the curriculum and would not be able to compete in the workforce or in higher education institutions. His answer to this dilemma was a set of standards for teachers and students to hold them accountable for curriculum (p13).

Heubert and Hauser (1999), members of the Committee on Appropriate Test Use for the National Research Council, stated that the basic principles for use of high-stakes tests are:

> The use of tests in decisions about student tracing, promotion, and graduation is intended to serve educational policy goals, such as setting high standards for student learning, raising student achievement levels, ensuring equal educational opportunity, fostering parental involvement in student learning, and increasing public support for the schools. (p. 2)

The public looks at high stakes tests to judge how well the public schools are doing and "policymakers see them as a way to raise standards and achievement and hold students and educators accountable" (Walker, 2000, p. 1).

Today many high stakes tests, like the Praxis (ETS the Praxis Series, 2011), and TOEFL (International Student Guide to the United States of America, n.d.), offer test-takers the choice of test modes. In addition to PPT, TOEFL dropped the CBT in favor of an internet based test (IBT) (TOEFL, 2011). Others such as the Medical College

Admission Test (MCAT) (Kaplan Test Prep, n.d.), the National Institute for Certification in Engineering Technologies (NICET) (NICET, 2008) and state standardized tests, such as the Florida FCAT (Florida Department of Education, "Frequently Asked Questions," n.d.), have changed from paper-pencil to computer-based administration. The Graduate Records Exam (GRE) computer version is taken in all areas of the world other than those where it is not available and only the PPT is offered (ETS GRE, n.d.). Assessments such as the American College Test (ACT) (ACT, 2010), and Scholastic Aptitude Test (SAT) (International Student Guide to the United States of America, n.d.), and Law School Admissions Test (LSAT) (International Student Guide to the United States of America, n.d.), continue to use paper-pencil application. When students, school districts, state education agencies, colleges, and the federal government put so much stock in standardized tests; many become concerned about the mode of administration. The Texas Education Agency (2008) stated that in a "high-stakes testing situation, schools may be reluctant to test students in a non-preferred mode" (p. 10). The investigators study looked at CGPA and the effect on preferred testing mode.

School districts, individual schools, and states may experience high stakes for their standardized tests. Follow up to these tests may include the following: examining student achievement, identifying effect on teachers and administrators, allocating resources, examining curriculum, and accounting to parents and the public. Therefore, the investigator felt it was important to analyze the effect of differences in test mode on these high-stakes tests.

End-of-Course Testing

Missouri's State Board of Education, in order to comply with the Outstanding Schools Act (Senate Bill 380), passed by the Missouri legislature in 1993, developed and implemented an assessment program to "measure student proficiency in the knowledge, skills, and competencies identified" (MODESE, 2009e, p. 7) in the Show-Me-Standards. These standards are to define "skills and competencies necessary for students to successfully advance through the public school system, prepare for post-secondary education and the workplace, and participate as citizens in a democratic society" (MODESE, 2009e, p. 7). Listed below are the purposes that would form the core of the Missouri Assessment Program.

- "Improving students' acquisition of important knowledge, skills, and competencies;
- Monitoring the performance of Missouri's educational system;
- Empowering students and their families to improve their educational prospects; and
- Supporting the teaching and learning process" (MODESE, 2009e, p. 7).

Complying with NCLB legislation, State leaders put in place new grade level assessments in the spring of 2006, and changed them further for the 2008-2009 school year. As a result of this last change the MAP assessments at the high school level would be replaced by EOC Assessments (MODESE, 2009e). "The EOC Assessments were created to address the needs of Missouri districts, schools, teachers, and students, while also meeting state and federal requirements" and several purposes were identified" (MODESE, 2009e,, p. 9). One purpose was to assess if a student was ready for a post-secondary institution. For instance had they learned the science curriculum if they were

preparing to go into premed. Additionally the EOC would reflect a student's strengths or weaknesses. This information could be used by school districts to evaluate their programs and identify areas for improvement in curriculum. The EOC tests and the materials they cover would communicate to students what they were expected to learn while in high school. Due to NCLB these assessments would meet the requirements for state and national accountability.

In the fall of 2009, the first Social Studies EOC test was added to the state's assessment program. Students in United States Government classes, across the state, were tested in fall, spring, and summer semesters, based on the semester the student was enrolled in a United States Government class. Initially all MAP tests were PPT administration and were the precursor of PPT EOC tests. During the 2009-2010 EOC test administration periods, districts were allowed to choose either CBT or PPT modes of administration. For the 2010-2011 school year DESE mandated that all EOC testing would be administered by computer.

Computer-Based Testing versus Paper-Pencil Testing

With the advent of choice of test mode administration organizations began to make comparisons in CBT and PPT in their search for testing practices to aid them in meeting NCLB's mandate. Florida's Department of Education, in their review of CBT, indicated that CBT first appeared in the early 1980s (Florida Department of Education, 2006). CBT may also be referred to in the literature as online testing, electronic testing, computerized testing, CBTs, or eTesting. Benefits to computerized testing include "more efficient test administration, flexible scheduling, quicker score reporting, more accurate examinee ability estimation, and expanded content/construct coverage" (Wan, Keng,

McClarty & Davis, 2009, p.1). Poggio, Glasnapp,Yang and Poggio (2005) found testing mode was changing to CBT in order to reduce score reporting time, provide continuous testing opportunities, improve security measures, reduce the cost of printing, handling, mailing and administering the test. Bodmann and Robinson (2004) found that CBT assessments were completed faster than PPT. *Education Week* published an article on testing, discussing the issue of cost and choice of test mode.

> Richard Swartz, a senior research director at the Educational Testing Service, in Princeton, N.J., (who) estimates that the actual costs of putting a test online and building a customized scoring model are comparable to those of developing a good paper-and-pencil exam . . . (but) once the tests are implemented, he adds, the difference in scoring costs is enormously in favor the computer.("Tech's Answer to Testing," 2003, para. 8)

Bodmann and Robinson (2004), in their research on CBTs versus PPTs, discussed the advantages of CBT over PPT and found the following advantages: easier to administer, easier to grade, faster tracking of grades, better standardized test conditions, easier to reduce cheating, and provides students the opportunity to choose when to take the test. Wang believed

> CBT delivery is gaining popularity over the traditional PPT delivery due to the several potential advantages that it offers, such as immediate scoring and reporting of results, more flexible test scheduling, the opportunity to include innovative item formats that are made possible by the use of technology, and reduced costs of test production, administration, and scoring. (Wang & Shin, 2009, p. 1)

Children preferred using a computer was a conclusion of Sim and Horton's (2005) study, but "the majority of the children performed better on paper than computer although there was no significant difference" (p. 3613). Studies have been done which indicate that with today's computer savvy generation, they are finding CBT's easier to take (McClarty et al, 2006). The National Institute for Certification in Engineering Technologies (NICET) is changing from PPT to CBT because of recent technology advances, best practices in testing, and the interest in increasing the value of the NICET certification (,NICET, 2008, para. 7). Russell and Plati (2000) found in their experiment "that students who wrote their compositions on computer produced longer responses that received higher scores" thus demonstrating CBT preference for state-mandated writing assessments (p. 26). Along with the abundant reasons for movement to CBT there is also considerable research that indicated PPT has less disadvantages than CBT.

With regard to some of the constraints for examinees who take computerized tests, the National Center for Fair and Open Testing (NCFOT, 2007) investigated the following issues: inability to underline or scratch out, greater length of time to read screens, greater difficulty finding errors on the screen, and difficulty in checking previous items. Mentioned in the GRE program are advantages to computer testing included taking the test at any time and taking it in a small less stressful venue; while disadvantages included inability to return to previous questions, it is easier to misread computer screens, and difficulty of going back and forth between computer screen and scratch paper (OneStopGRE, 2009). Mayes, Sims and Koonce (2001) stated "those who read from a VDT (video display terminal) took significantly longer than those reading from paper," (p. 1). Puhan, Boughton , and Kim (2007) felt that CBTs were more difficult

than the PPT test version. By 1993 Mead and Drasgow had conducted a study which found PPT scores to be slightly higher (as cited in McClarty & Davis , 2006, p.4). The Florida Department of Education (2006) found that "while some early studies suggested that students who had less experience with computers would score lower on computer-administered tests, recent studies find no evidence of such a disadvantage" (p. 3). Dillon, McKnight, and Richardson (1988) felt that "although reading from computer screens may be slower and sometimes less accurate than reading from paper, no one variable is likely to be responsible for this difference" (Section 3.12). Some of the variables they reviewed were speed, accuracy, fatigue, comprehension, preference, orientation, eye movement, visual angle, ratio of width to height, display characteristics, and user characteristics (Dillon et al., 1988). "Most studies showed higher scores for paper-and-pencil exams, but a few have found advantages for those who take computerized tests" (NCFOT, 2007, p. 1).

Many school districts lack the necessary infrastructure or technology equipment to have their students test on computers; therefore they use both CBT and PPT modes of administration. Awareness of issues, both positively and negatively affecting both modes of administration, was important to states as they adopted new testing requirements. Observing some problems with test mode, National Center for Fair and Open Testing (NCFOT, 2007) identified the following: rushing new tests into operation without providing evidence of comparability, test-maker claims that are not supported, studies showing higher scores for PPT, and test questions that might perform differently on each type of test. Shead (2006) discussed another issue contemplated when choosing CBT or PPT; the differences in what part of the brain is triggered when writing with

paper-pencil or computer and that "the conclusion of researchers was that we think significantly differently when writing by hand than we do when using a computer"(para.2). Bodmann and Robinson (2004) found the only test mode effect was "a difference in completion time between PPTs and CBTs" and that reviewing and changing answers accounted for this difference (p. 57). An issue that crops up regularly is whether the two scores, CBT and PPT, can be compared.

There is some concern in the research that comparing the scores from the two modes of administration may not reflect the examinees proficiency in the subject matter, but rather how proficient they are in computer usage (Puhan et al, 2007). When the choice of mode of testing is left up to the examinee several issues come in to play. Bernt, Bugbee, and Arceo (1990) found no relationship between computer experience and test mode preference, the more computer usage the individual had the less negative they were towards using computers, and examinees beliefs about benefits of computer testing were related to their test mode preference.

When comparing CBTand PPT, concerns also exist in regard to subgroups taking the tests. Gallagher, Bridgeman, and Cahalan (2002) examined data from testing programs such as GRE, SAT, Praxis, TOEFL, and GMAT, with regard to gender and racial-ethnic subgroups. They concluded: African American and Hispanic examinees benefitted slightly from CBT, while women performed better on PPT. In regard to subgroup testing, Wallace and Clariana (2005) felt "the performance gap which already exists on multiple-choice tests between men and women, ethnic groups, and persons from different socioeconomic backgrounds could widen as a result of computerized testing" (p. 172). They also ascertained that with gender differences "performance on the computer-

administered tests was significantly greater than performance on the identical paper-based tests" (Wallace & Clariana, 2005, p. 176) for both genders. "Females generally scored lower than males under both computer- and paper-based test administrations on the test given early in the course (the DCC test), but then scored highest on the computer-administered Final examination" (Wallace & Clariana, 2005, p. 177).

The Florida Department of Education ("What Do We Know About Choosing to Take a High-stakes Test on a computer?", 2006) in their comparability review, looked at 97 cases reviewed by others, and found that 74 of these studies discovered the two modes of testing to be comparable, with eight saying CBTs were more difficult and 15 designating PPT more difficult. While their study examined CGPA and test mode administration for the NCLB subgroups, they also realized that "most studies do not focus on comparability for different subgroups of students" (Florida Department of Education, 2006, p. 3). McClarty and Davis (2006) discovered in their review of comparability studies, that in earlier work the two test modes were not comparable, and showed a favoring for PPT, whereas since 1993 "test scores tended to be higher for those testing on paper-and-pencil rather than computer, the magnitude of the difference was extremely small" (p. 4). In her review of research on comparability of test modes, Paek (2005) concluded "The K-12 comparability studies to date show that, in general, computer and paper versions of traditional multiple-choice tests are comparable across grades and academic subjects" (p. 17). A study conducted by Millsap (2000) "found no significant difference between test administration modes" and "concluded that computer-administering tests identical to those typically administered in the traditional paper and pencil manner had no significant effect on achievement" (p. 58).

Considerable research can be found to support both sides of the question regarding comparability of CBT versus PPT administration. Leeson (2006) investigated differences that can occur between the two types of test administration and found factors such as, demographics, gender, ethnicity, cognitive processing, computer interaction anxiety, and familiarity with computers can limit the value of comparing the tests. Clariana and Wallace (2002) demonstrated that gender, computer familiarity, and competitiveness were not related to test mode, while content familiarity was related to test mode. Much inconsistency existed in research comparing the two test modes, and many researchers felt more testing on comparability needs to be conducted, particularly as society becomes more highly computer oriented. Because these standardized tests carry such high stakes as graduation, entry into various programs and institutions, scholarships, and allocation of funding; it is necessary to be sure favoritism is not shown to one group of test takers over another.

Differentiation in the Classroom in Regard to Test Administration

Indications from the Florida Department of Education (2006) are that sometimes it is better to let either students or teachers decide the test mode.

> Prior to moving to computer-based testing, the Department studied whether there is a difference in performance between assessments taken on the computer and on paper … At that time, the preponderance of studies of the comparability of K–12 computer- and paper-administered multiple-choice tests showed differences that were either statistically not significant or of no practical significance. However, other studies have shown advantage to either paper or computer administration. Each year, more studies are being conducted, and our understanding of potential

differences in the testing modes will continue to increase. ("Florida End of Course," Question 10)

Hall et al. (2003) discussed differentiation of instruction as a need to look at a students' variety of background knowledge and decide the instruction they need based on their abilities. They indicate the purpose is to come up with a process that allows students of differing abilities in the same class to continue to grow and succeed according to their needs. The State of Florida examined the difference between PPT and CBT before they moved to CBT and found that most of the comparability studies for K-12 showed differences that had no practical or statistical significance and that in the future more studies were being conducted that would enhance our understanding of possible differences in testing modes (Florida Department of Education, 2006, "Florida End of Course," Question 10).

Research has also been completed comparing CBT and PPT for a particular curriculum. Russell and Plati (2000) looked at Language Arts testing and concluded that when state testing programs used open-ended questions, students should be provided "the option of composing responses on paper or on computer" (p. 34). In a study examining test scores of students in an Intermediate Accounting I class, where all students participated in the same class curriculum, Maguire, Smith, Brallier, and Palm (2010) reported "results indicated that students who completed all assessments electronically scored significantly higher than those students completing all assessments via pencil and paper" (p. 1) and "CBT resulted in a higher average score than the traditional method" (p. 3). Specifically searching for research that demonstrated student choice of test mode, whether for classroom or high-stakes testing, produced no results.

Brain Function Related to Reading On-line Versus on Paper

In the area of brain function related to reading Gray (2010) indicated "Research shows that the brain functions differently when reading online versus reading a book, and different formats complement different learning styles" (p. 30). A great deal of research existed on differentiation in the classroom and being aware of the variety of learning styles students bring to the table. If, as Gray claims reading online or on paper will complement different learning styles, then perhaps students should take tests that match their learning style.

With regard to score comparability of assessments**,** although availability of technology is increasing there are still states which, like Missouri in previous school years, had some schools take the test paper-pencil, while others were using computers. The Texas Education Agency (2008) noted that "many schools do not have the infrastructure and equipment to test all of their students by computer" (p. 6). With states using both modes of administration and comparing grades across the state, many believe there is a need to demonstrate the two modes of testing are comparable. McClarty and Davis (2006) in their research, discussed the need to conduct comparability studies during this transition period (p. 3). "Evaluating test comparability is … essential so that no student is disadvantaged by taking a test on the computer or on paper" (McClarty & Davis, 2006, p.19). The investigators study used statistics for the Missouri transition period from PPT to CBT.

In order for a state to use the scores to compare districts across the state and to evaluate whether NCLBs standards are being met, there should be evidence that PPT and CBT modes of administration and exam scores are comparable (Texas Education

Agency, 2008). Paek (2005) stated not to take for granted the two testing modes are comparable, but to explore all the differential effects between CBTs and PPTs. The U.S. Department of Education (2009) NCLB peer review cited "If the State administers both an online and paper and pencil test, has the State documented the comparability of the electronic and paper forms of the test?" (p. 46). State accountability for achievement testing in K-12, brought about by NCLB, have initiated questions about the validity of test scores used to compare institutions (Wise, Kingsbury, Thomason & Kong, 2004). The Texas Education Agency (2008) reported "whenever paper-pencil and computer-based assessments of the same content are administered, professional testing standards and federal accountability both require evidence showing comparability of test scores obtained in the two administration modes" (p. 2). The American Psychological Association (APA) Professional Test Standards and the Joint Standards for Educational and Psychological Testing demonstrate the need to study comparability of scores between the two test modes (Texas Education Agency, 2008). Further stated was "most comparability studies conducted across a number of different state testing programs have found test scores to be comparable across the two administration modes" (Texas Education Agency, 2008, p. 2). However, a warning was given that there could be differences from one content area to another . When the GRE board decided to switch from PPT to CBT, they ran a field test in 1991 to test for comparability and with their results considered testing mode scores to be comparable (Schaeffer, Bridgeman, Golub-Smith, Lewis, Potenza & Steffen, 1998). Way (2006) stated "professional testing standards require that states provide evidence about score comparability when assessments are delivered both online and by paper" (p. 2)

CBT/McGraw-Hill (2003) explained the importance of score interchangeability and what is required for "test scores to be considered interchangeable between paper-and-pencil (p&p) and computer-based or on-line modes of test administration" (p. 1). Listed in their article were a number of standards or guidelines that should be followed to address this comparability between CBT and PPT, which include:

- "International Guidelines on Computer-Based and Internet Delivered Testing (Draft Version 2003): International Testing Commission A Code of Practice for the Use of Information Technology for the Delivery of Assessments (2002: British Standards Institution
- Guidelines for Computer-Based Testing (2002: Association of Test Publishers)
- Guidelines for the Development and Use of Computer-Based Assessments (2002: British Psychological Society)
- Standards for Educational and Psychological Testing (1999: American Educational Research Association/American Psychological Association {AP}/National Council on Measurement in Education})
- Guidelines for Computer-Based Tests and Interpretations" (1986, APA) (as cited inCBT/McGraw-Hill, 2003, p. 2).

Indications were that with NCLB legislation in place, agencies may be switching to on-line testing (CBT/McGraw-Hill, 2003). "Interchangeability is required when students may take the same test in either mode" (CBT'McGraw-Hill, 2003, p. 1).

As Missouri gave both CBT and PPT EOC exams during the same school year, the investigator contacted MODESE regarding comparability studies. After contacting

VanDeZande, Director of Assessment, Office of College and Career Readiness, MODESE, and inquiring as to whether Missouri had done a comparability study for their EOC tests regarding PPT or CBT administration; the investigator was informed no such study was done for the United States Government EOC exam. As of November 2010 MODESE was not planning a separate validity/reliability study for the United States Government EOC exam, the test used in this study (VanDeZande, personal communication, November 12, 2010). In order to comply with NCLB, and on behalf of the state of Missouri, Riverside Publishing conducted a PPT versus CBT comparability study for the English II, Algebra I, and Biology tests during the 2008-2009 school year. VanDeZande forwarded the 2009 copy of this report. MODESE (2009f) stated

> The comparability of a computer-based assessment to its paper-and-pencil counterpart cannot be assumed. Conceivably the mode of administration may affect the difficulty of the test, either through an overall shift in difficulty or through an item-b-mode interaction. Riverside Publishing conducted the current study for the purpose of describing a strategy for evaluating the comparability of Missouri's P/P and online EOC Assessments and to provide a summary of several analyses performed to determine the comparability of the two modes for the spring 2009 administration. (p. 1).

Their evidence suggested

> there is little appreciable difference in the factor structures of the tests delivered by the different modes, 2) there is little evidence of DIF to indicate that certain

> types of items function differently by mode; and 3) there do seem to be differences in mean performance by mode for the Algebra and English II tests, but the differences seem to be practically small. (MODESE, 2009f, p. 27)

Kim and Huynh (2007) studied Algebra and Biology EOC results from school districts in a southeastern state and discovered "some empirical evidence of comparability of statewide PPT and CBT in Algebra and Biology at the item-level, subtest-level, and whole test-level" (p. 25). A comparison of PPT versus CBT of student performance on a statewide EOC English test led Kim and Huynh (2007) to conclude the students' scores for both modes of testing were comparable. Paek (2005) deduced "that the computer may be used to administer tests in many traditional multiple-choice test settings without any significant effect on student performance" (Abstract, para. 3). When conducting a study of two business information systems courses at the university level Bartlett, Alexander, and Ouwenga (2001) concluded their study provided "evidence that online testing provides results that are equivalent to traditional paper and pencil testing in relationship to student test scores" (p. 5). Kapes, Martinez, Chui-Fung, Slivinski, and Hardwick, (1998) looked at 11th and 12 th grade students occupational competency tests and determined that CBT and PPT versions of the test were equivalent. Choi and Tinkler (2002) studied third and tenth graders tested in reading and math and their results indicated "identical items administered in different modes were generally more difficult on the computer" (p. 8). CBTs may favor students with greater computer experience (Choi & Tinkler, 2002, p. 10). Al-Amri (2008), when studying the comparability of CBT vs. PPT discovered "testing mode has almost no significant effect on the overall validity and reliability of the tests" (p. 41).

The NCFOT (2007) reported "test-makers claims that the scores of computerized and pencil-and-paper tests are equivalent are inadequately supported. In fact, research studies find there usually is a difference" (Unresolved Problems, para. 1). After reviewing comparability studies, Pommerich (2004) found there are mixed results, but indications were the "more complicated it is to present or take the test on computer, the greater the possibility of mode effects" (p. 3). Wang and Shin (2009) summed up the importance of comparability stating "The comparability between the alternative test versions cannot be taken for granted and related investigations have to be done to ensure that the examinees are not treated unfairly due to the testing mode" (p. 5).

Relationship of CGPA to Computer-based or Paper-pencil Test Administration

A 2001 study by members of the Ball State University College of Business employed an investigation where both test groups had a 2.5 grade point average. Comparisons of the two groups provided evidence that computer-based and paper-pencil test results were equivalent (Bartlett, 2001). Ball State University's study did not look at individual CGPAs, but looked at the group average. The state of Florida's literature review on taking high stakes tests on computers found comparability studies for computer experiences, race/ethnicity, gender, and demographic subgroups; but no information with regard to GPA (Florida Department of Education, 2006). This information was being provided to schools in their state to help them decide whether to take paper or computerized versions of their standardized tests. Data from the investigators study will be forwarded to MODESE in order to provide additional research that may help when they review modes of test administration to be used for Missouri's EOC testing.

Summary

As evidenced from this literature review, there is very little information available regarding the correlation between CGPA and test mode; therefore it is the intention this study will help to fill this research gap. A myriad of research exists on the comparability of CBT and PPT exams, and the above is just a small slice of what is available. There are many variables including student cognitive ability, computer skills, proficiency in subject matter, examinee motivation, the variety of computer models utilized, and availability of technology, which can influence the outcomes of comparability studies.

There are so many variables that dictate whether CBT or PPT would be a better choice for the test taker, and perhaps in light of the educational best practice of differentiation in the classroom, we need to assess each student's aptitude for computer usage and have them test in the mode most conducive to their testing style. Starting with Chapter 3, the following chapters will discuss this study's methodology and results with regard to the influence of CGPA on test mode administration.

Chapter 3: Methodology

This quantitative study analyzed the causal-comparative correlation between two variables: CGPA, and state mandated EOC, United States Government exams, through either computer or paper-pencil administration. As required by the state of Missouri, during the course of United States Government classes, students are required to take this multiple choice assessment. The purposes of this study were to determine if CGPAs were a predictor of student success on computer-based and paper-pencil exams, and to possibly provide information to the MODESE, and school districts, as to the benefits of one type of test administration over the other or the benefits for providing both options. The rationale for the project was to determine which approach to testing, CBT, or PPT, would provide the most accurate information of student mastery for school and state accountability. In this study, School District A administered computer-based tests, while School District B utilized paper-pencil testing.

NCLB required states to implement statewide accountability systems covering all public schools and students. These systems must be based on challenging state standards in reading and mathematics, annual testing for all students in Grades 3-8, and annual statewide progress objectives ensuring that all groups of students reach proficiency within 12 years. Missouri requires EOCs for Grades 9 and above, to meet AYP.

> Assessment results and State progress objectives must be broken out by poverty, race, ethnicity, disability, and limited English proficiency to ensure that no group is left behind. School districts and schools that fail to make Adequate Yearly Progress (AYP) toward statewide proficiency goals will, over time, be subject to improvement, corrective action, and restructuring measures aimed at getting them

back on course to meet State goals (U.S. Department of Education, 2004, Increased Accountability, p. 1).

The federal government does not provide standardized tests to meet these professional goals, but leaves the choice of these assessments to the individual states to provide their own system of assessments; therefore Missouri has EOC tests in the areas of Science, Math, English and Social Studies. With the help of Riverside publishing and teams of teachers, MODESE prepared and provided common end-of-course tests for each subject area, in both paper-pencil and computer-based formats for the 2009-2010 school year. The MODESE stated the following are the "purposes for the Missouri End-of-Course (EOC) Assessments:

- "Measuring and reflecting student's mastery toward post-secondary readiness.
- Identifying students' strengths and weaknesses.
- Communicating expectations for all students.
- Serving as the basis for state and national accountability plans.
- Evaluating programs" (MODESE, 2009a, p. 1).

The Parkway School District felt "End-of-course exams will provide a valid and reliable method for assessing students' knowledge of Missouri's Course-Level Expectations (CLEs). They will also allow classroom teachers to incorporate statewide assessment results into students' course grades" (Parkway School District, 2010, p.1)

Questions Addressed in the Study

The following questions were addressed in the study:

1. What is the relationship between students' cumulative grade point averages and their exam scores for United States Government End-of -Course exams, through

paper-pencil or computer-based administration?

2. What is the relationship been students' cumulative grade point averages and their exam scores for United States Government End-of-Course exams, through paper-pencil or computer-based administration, disaggregated by No Child Left Behind subgroups?

3. Should the Missouri Department of Elementary and Secondary Education continue with their mandate for computer-based administration of End-of-Course exams, or would choice of mode of administration be in the best interest of students; better demonstrating student mastery of the curriculum as determined by their performance on these exams, and therefore benefit their district and the state?

Independent Variable

United States Government EOC exam scores, attained through CBT or PPT, administered during the 2009-2010 school year, was used as the independent variable in this study.

Dependent Variable

The dependent variable was high school CGPA, measured at the end of the 2009 2010 school year.

Hypotheses:

Null Hypothesis 1: For students taking the United States Government End-of-course examination by paper-pencil mode of delivery, compared to students taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government End-of-course exam.

Null Hypothesis 2: For students taking the United States Government End-of-course examination by paper-pencil mode of delivery, there is no relationship between student cumulative grade point average and United States Government End-of-course exam score.

Null Hypothesis 3: For students taking the United States Government End-of-course examination by computer-based mode of delivery, there is no relationship between student cumulative grade point average and United States Government End-of-course exam score.

Null Hypothesis 4: For students taking the United States Government End-of-course examination by paper-pencil mode of delivery, students with a cumulative grade point average of 2.5 or higher will not score higher than students with a cumulative grade point average of less than 2.5, on the United States Government End-of-course exam.

Null Hypothesis 5: For students taking the United States Government End-of-course examination by computer-based mode of delivery, students with a cumulative grade point average below 2.5 will not score higher than students with a cumulative grade point average of 2.5 or above, on the United States Government End-of-course exam.

Population

The population studied included all students who took the United States Government EOC test in two St. Louis County suburban school districts for the 2009-2010 school year. For the purpose of this study they will be designated School District A and School District B. Applications to perform research were submitted and accepted by both school districts. Data was made available September 2010 for the 2009-2010 school year. School District A administered the test via computers, while School District

B used a paper-pencil administration. The Directors of Data Analysis from each district provided an electronic database of all 2009-2010 exam scores, correlated with the students' cumulative CGPA, as of the end of the 2009-2010 school year. Data was further disaggregated by the following categories: gender, score, percent correct, raw score multiple choice (MC), total raw score, gifted, individualized education program (IEP), English as a second language (ESL), ethnicity, and free and reduced lunch (FRL).

School District A and B Demographics

The researcher chose School District A and B because of their similar demographics and size, as noted in Table 1 and 2.

Table 1

School Districts A and B Demographics

	School District A	School District B
Total Population within district boundaries	140,660	155,596
Total Student Population	18,301	22,318
Attendance Rate	95%	95.4%
Graduation Rate	93.8%	95.2%
Dropout Rate	1.4%	1.2%
Elementary Schools	18	19
Middle Schools	5	6
High Schools	4	4
Non-traditional High School	1	1
Early Childhood Center	1	1
Average Household Income	$108,844	$97,496*
Total Number Employees	2,967	3,402
Certified Classroom Teachers	99.7%	99.7%
Certified Teachers with Advanced Degrees	68.1%	68.2%

*2000 Census

Table 2

School Districts A and B Racial Demographics

	School District A	School District B
Racial Makeup of District * **		
Caucasian	89.4%	94.6%
Asian	5.8%	2.3%
African-American	2.8%	1.3%
Other	2%	1.8%
Racial Makeup of Student Body **		
Caucasian	70.4%	82.3%
Asian	11%	5.1%
African-American	16.1%	10.3%
Other	2.5%	2.0%

* 2000 Census **Statistics are from MODESE website.

Remaining statistics were found on either School District A or School District B home websites. School Districts A and B are located in the same geographic area of West County in St. Louis, Missouri. After looking at the similarities in the statistics above, it was felt the two districts were comparable and provided a commensurable source of testing data.

Testing Procedures

The MODESE dictated all government students in Missouri would take the Government EOC assessment during the semester they were enrolled in this course. MODESE provided a time frame (usually a two week period) for each semester, during which schools were required to test the students. Throughout the state students tested during the same time period. During the 2009-2010 school year, each district had the choice of CBT or PPT administration. In this study one school district used CBT, while

the other used PPT. MODESE graded the tests and scores were then returned to the school districts.

Data Analysis Procedures

In order to discuss the relationship between CGPA and mode of test administration, information was collected to describe the study. "A variable is a characteristic or event that can assume different values" (Bluman, 2008, p. 3). Data used in this study were the values that the variables assumed. Using the methods of inferential statistics, this study will attempt to generalize from sample School Districts A and B, to general populations. Hypotheses will be tested to determine the relationship between CGPA and test mode, the quantitative variables, and to make predictions concerning student success on either CBT or PPT.

Due to the fact the data obtained from the two school districts numbered over 5,000 pieces, the sampling method used for this study was the Random method, whose subjects were selected by random numbers generated through a computer web based service, Research Randomizer. Fifty students from each of the two school districts were selected for inclusion in this study.

Table 3

Demographic Data for Random Samples for School Districts A & B

	Male	Female	Have IEP	No IEP	Asian	Black	White	Hispanic	FRL
District A	23	27	8	42	6	8	34	1	10
District B	26	24	0	50	2	9	38	1	5

Note: This table demonstrates a frequency count. FRL= Free and Reduced Lunch: IEP = Individualized Education Program

Subjects in four NCLB subgroups including gender, ethnicity, FRL, and disabilities, were also selected by additional random sampling, in order to compare type

of testing, CBT or PPT, in relation to student CGPA.

IBM's Statistical Package for Social Sciences (SPSS) software was used to manipulate the data accumulated for this study. SPSS is a computer program that is widely used in the Social Studies field for statistical analysis (University of South Florida, 2009, para. 4).

Bluman (2008) reported that in simple correlation studies, "the researcher collects data on two numerical or quantitative variables to see whether a relationship exists between the variables" (p.523). Box plots were used to determine if a relationship existed between the two variables studied, CGPA and method of test administration. A statistical t-test was conducted to test for difference in the mean, and to decide whether to reject or not reject the null hypotheses.

Because some samples were small in size (n<15), results of the *t*-test for difference in means, comparison of samples with GPA 2.5 and above to samples with GPA below 2.0, were checked with the more conservative Chi Square Test for Homogeneity and Confidence Interval Test for Difference in means. The smallest sample size allowable for this calculation was 5. When results contradicted those yielded with the *t*-test, the results from the Chi Square Test for Homogeneity were recorded.

Confidence intervals were calculated to provide a secondary test to support results of the *t*-test for difference in means for small samples ($n < 15$). Minimum requirements for sample size were calculated and considered. When results were contradictory, they were reported in descriptive format only.

Correlation

The purposes of this study are to answer through the use of statistics the following

questions:

* Are CGPA and test administration mode related?

* If so, what is the strength of the relationship?

* What type of relationship exists?

* What kind of prediction can be made related to who should take CBT vs PPT, based on their CGPA?

The first two questions will be answered by testing appropriate null hypothesis statements through use of calculation of the Pearson Product Moment Correlation Coefficient (PPMCC), which "measures the strength and direction of the linear relationship between the two variables" (Bluman, 2008, p. 525). Answering the question as to what type of relationship exists; the study will look at a simple regression. Based on the strength of the relationship, the researcher may be able to make mild predictions about best method of testing for students.

Descriptive Statistics

Tables 4 through 7 include the descriptive statistics for either School A or School B with CGPA either below 2.5 or 2.5 and higher.

Table 4

School A EOC Scores, CBT, CGPA Below 2.5

	N	Minimum	Maximum	Mean	Std. Deviation
School A EOC Scores	12	154.00	212.00	184.0833	16.41761
Valid N (listwise)	12				

Table 5

School A EOC Scores, CBT, CGPA 2.5 and Above

	N	Minimum	Maximum	Mean	Std. Deviation
School A EOC Scores	38	165.00	250.00	220.7368	20.12083
Valid N (listwise)	38				

Table 6

School B EOC Scores, PPT, CGPA Below 2.5

	N	Minimum	Maximum	Mean	Std. Deviation
School B EOC Scores	14	167.00	228.00	202.5000	14.93962
Valid N (listwise)	14				

Table 7

School B EOC Scores, PPT, CGPA 2.5 and Above

	N	Minimum	Maximum	Mean	Std. Deviation
School B EOC Scores	36	194.00	250.00	223.8333	13.58676
Valid N (listwise)	36				

Hypothesis Statistical Analysis

Null Hypothesis 1: For students taking the United States Government EOC examination by paper-pencil mode of delivery, compared to students taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam. In order to test Null Hypothesis 1, as the standard deviation is unknown, a *t*-test for difference in means was conducted.

Null Hypothesis 2: For students taking the United States Government EOC

examination by paper-pencil mode of delivery, there is no relationship between student cumulative GPA and United States Government EOC exam score. To determine the strength of the relationship between CGPA and EOC exam scores, the researcher calculated the Pearson Product Moment Correlation Coefficient (PPMC) and check for significance of the value at the alpha level of 0.05.

Null Hypothesis 3: For students taking the United States Government EOC examination by computer-based mode of delivery, there is no relationship between student cumulative GPA and United States Government EOC exam score. As described for Null Hypothesis 2, the researcher tested the strength of the relationship between CGPA and EOC exam scores using the PPMC.

Null Hypothesis 4: For students taking the United States Government EOC examination by paper-pencil mode of delivery, students with a cumulative GPA of 2.5 or higher will not score higher than students with a cumulative GPA of less than 2.5, on the United States Government EOC exam. A *t*-test for difference in means was conducted to test Null Hypothesis 4, with an alpha level of 0.05.

Null Hypothesis 5: For students taking the United States Government EOC examination by computer-based mode of delivery, students with a cumulative GPA below 2.5 will not score higher than students with a cumulative GPA of 2.5 or above, on the United States Government EOC exam. As in Null Hypothesis 4, for the PPT mode of administration, for Null Hypothesis 5, the CBT mode of administration, the researcher again conducted a *t*-test for the difference in means, with an alpha level of 0.05.

Threats to Internal and External Validity

The United States Government EOC test, used in this study, was developed by

Riverside Publishing Company for the MODESE. In the Missouri End of Course Assessments Technical Report Phase II Assessments 2009-2010, Riverside Publishing demonstrated in Chapter 10 "evidence that scores from the Missouri End-of-Course (EOC) Assessments measure student achievement in a reliable manner and that the size of the measurement error associated with reported test scores is reasonable, especially at the Proficient cut score" (MODESE, 2010b, p. 181). Further, in Chapter 11, Riverside analyzed the validity of the EOC test with regard to how adequately and how appropriately the assessment measured proficiency of the Missouri content standards (MODESE, 2010b, p. 197).

This researcher's study did not take into account other variables that could affect the students' score on the test. Just a few of these include computer proficiency, testing comfort levels, reading skills, typing skills, differences in testing conditions, computer anxiety, cognitive processing, characteristics of computers being used, and test mode preference.

Limitations

This study was conducted using data from two large suburban high schools; therefore results may be biased due to the demographics of these districts. The demographics lean heavily towards white, middle to upper class students. Due to the use of these two particular demographic areas, the use of technology is more prevalent than might be found in other districts. As this study is comparing computer usage to paper-pencil usage, results may not be transferable when applied to other school districts. For instance the average household income was around $100,000 and around 90% of the

population was Caucasian. School districts of lower socio-economic status and racial diversity may find the results of this study to be nontransferable.

Summary

This study investigated the relationship between students' CGPA and their test scores on EOC, United States Government exams, considered a high-stakes test, and the manner in which the test was administered. Two large suburban school districts supplied the data for this quantitative study. If a relationship exists between CGPA, types of test administration, and resulting exam scores, then perhaps high stakes tests should be offered to students with a choice of formats. Or perhaps the students should be required to take the type of test, which their CGPA indicated would be best, in order to achieve their maximum scores.

In Chapter 4 the author discusses the statistical analysis of this study and interprets the data, in order to answer the questions of relationship between CGPA and modes of test administration.

Chapter 4: Results

This study analyzed the relationship between EOC test scores in United States Government classes given by PPT or CBT mode of administration and students' CGPA, at the end of the 2009-2010 school year. The purposes of this study are twofold. One was to determine if CGPAs are a predictor of student success on computer-based and paper-pencil exams. The second purpose of the study was to provide information to the Missouri Department of Elementary and Secondary Education (MODESE) and school districts, as to the benefits of one type of test administration over the other; or the benefits for providing access to both options.

Reliability/Validity of Government EOC Tests

Riverside Publishing conducted a reliability and validity study for the English II, Algebra I, and Biology EOC tests during the 2008-2009 school year. MODESE (2009f) reported that between the two testing modes, PPT and CBT, for these EOC tests there was little discernible difference. There has been, and are, no reliability and validity studies planned for the Government EOC test (J. VanDeZande, personal communication, November 18, 2010).

Data Collection

Directors of Data Analysis from two large St. Louis suburban school districts, one of which administered the test by paper-pencil, and the other by computer, provided an electronic database of all district 2009-2010 U.S. Government End-of-Course exam scores, matched with the students' cumulative CGPA, as of the end of the 2009-2010 school year. No student names or numbers were included with the data in order to protect student identity. Both school districts administered the test during the students'

junior year of high school. High school EOC test scores for United States Government exams were used as the independent variable in the study. Over 5,000 pieces of data were received from the two school districts. The Random Systematic sampling method was used to select 50 subjects from each school district for inclusion in the researchers study.

Table 8

EOC Results Matched with CGPA, Scores Below 2.5

School District A		School District B	
EOC Score	CGPA	EOC Score	CGPA
165	1.182	192	1.333
203	1.432	167	1.404
184	1.651	205	1.523
200	1.698	201	1.791
175	1.769	202	1.833
187	1.889	194	1.978
154	1.897	194	1.978
175	2.025	205	2.000
192	2.18	220	2.095
212	2.300	205	2.195
186	2.345	208	2.250
176	2.417	194	2.282
		220	2.410
		228	2.46

Table 9

EOC Results Matched with CGPA, Scores 2.5 and Above

School District A		School District B	
EOC Score	CGPA	EOC Score	CGPA
225	2.583	229	2.561
194	2.738	218	2.714
228	2.816	215	2.825
188	2.881	233	2.925
210	2.905	239	2.940
196	2.929	215	2.974
165	2.962	202	2.979
247	2.978	221	3.119
209	3.000	216	3.154
205	3.079	220	3.179
216	3.313	205	3.200
189	3.452	215	3.263
250	3.456	225	3.366
233	3.500	225	3.395
233	3.524	239	3.436
228	3.579	215	3.475
208	3.583	210	3.525
198	3.619	228	3.550
228	3.643	225	3.571
228	3.750	203	3.591
218	3.750	212	3.634
205	3.766	229	3.762
233	3.810	234	3.810
218	3.833	234	3.829
234	3.837	225	3.875
225	3.857	207	3.925
247	3.922	240	3.925
233	3.975	233	3.974
216	4.024	240	4.024
240	4.054	250	4.026
215	4.070	247	4.130
228	4.116	221	4.146
250	4.375	229	4.250
247	4.476	247	4.275
240	4.634		
250	4.667		

The dependent variable was CGPA, as measured at the end of the 2009-2010 school

year. Working with the hypothesis that students with a 2.5 or higher CGPA would score higher on PPT, while those with less than 2.5 CGPA would score higher on CBT; the researcher sought to establish a correlation between CGPA and mode of test administration.

Overall Data

Statistical Analysis

Statistical Package for Social Studies (SPSS) software was used to procure descriptive statistics, *t*-tests, Pearson Product Moment Correlation Coefficients, and to generate box plots for each hypothesis. SPSS is a computer program that is widely used in the Social Studies field for statistical analysis. (University of South Florida, 2009, para. 4) In order to test Null Hypothesis 1, as the population standard deviation is unknown, a *t*-test for difference in means was conducted.

Because some samples were small in size (n<15), results of the *t*-test for difference in means, comparison of samples with GPA 2.5 and above to samples with GPA below 2.0, were checked with the more conservative Chi Square Test for Homogeneity and Confidence Interval Test for Difference in means. The smallest sample size allowable for this calculation was 5. When results contradicted those yielded with the *t*-test, the results from the Chi Square Test for Homogeneity were recorded.

Confidence intervals were calculated to provide a secondary test to support results of the *t*-test for difference in means for small samples ($n < 15$). Minimum requirements for sample size were calculated and considered. When results were contradictory, they were reported in descriptive format only.

Null Hypothesis 1: For students taking the United States Government EOC

examination by paper-pencil mode of delivery, compared to students taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam.

According to the descriptive statistics in Table 10 there is no observable difference, which was supported statistically by running the *t*-test for difference in means [*t* (98) = 1.395, alpha = 0.05, *p* = 0.166]: therefore, the researcher did not reject the null hypothesis. When student scores were averaged there was no difference between those taking the test PPT and those taking the test CBT.

Table 10

Descriptive Statistics Comparing Schools A and B

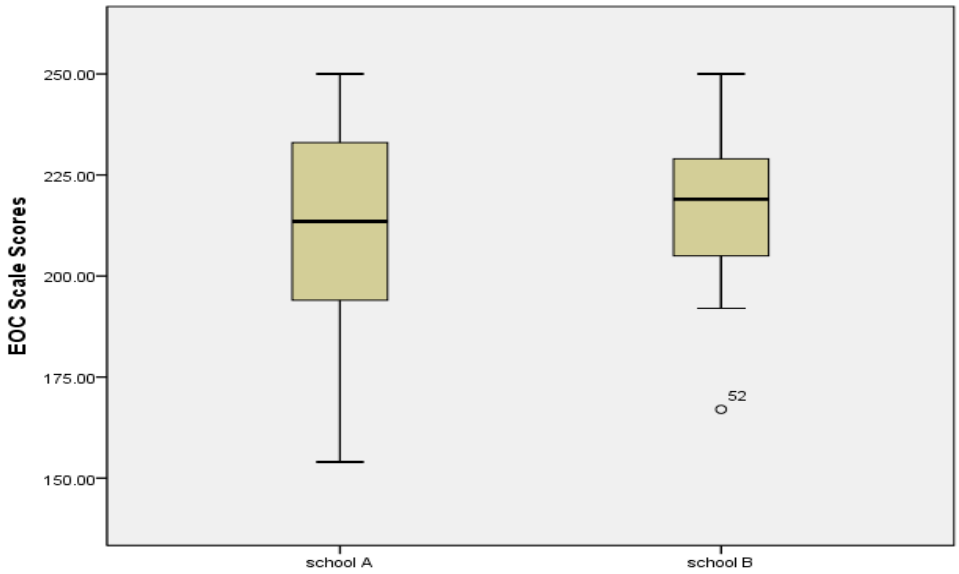
	School A & B	N	Mean	Std. Deviation	Std. Error Mean
EOC	school A	50	211.9400	24.82462	3.51073
	school B	50	217.8600	16.87289	2.38619

Note: School A used computer-based and School B used paper-pencil-based testing.

The box plot in figure one demonstrates the means for scores in school A (CBT) and B (PPT) are similar, but the ranges are drastically different. Students taking the test CBT had a wider range of scores than students taking the test PPT.

Figure 1- Box Plot, Comparison of EOC Scale Scores by School District

School District A = CBT School District B = PPT



Null Hypothesis 2: For students taking the United States Government EOC examination by paper-pencil mode of delivery, there is no relationship between student cumulative GPA and United States Government EOC exam score.

The Pearson Product Moment Correlation Coefficient was calculated using the data to check on the status of Null Hypothesis 2. According to Table 11, the Pearson Product Moment Correlation Coefficient demonstrates a strong correlation that is statistically significant [R (48) = 0.719, alpha = 0.05, p <0.0005] therefore the researcher rejected the null hypothesis and data supports the alternative hypothesis, indicating there is a strong positive relationship between students CGPA and EOC scores for students taking the PPT form of the U.S. Government EOC.

Table 11

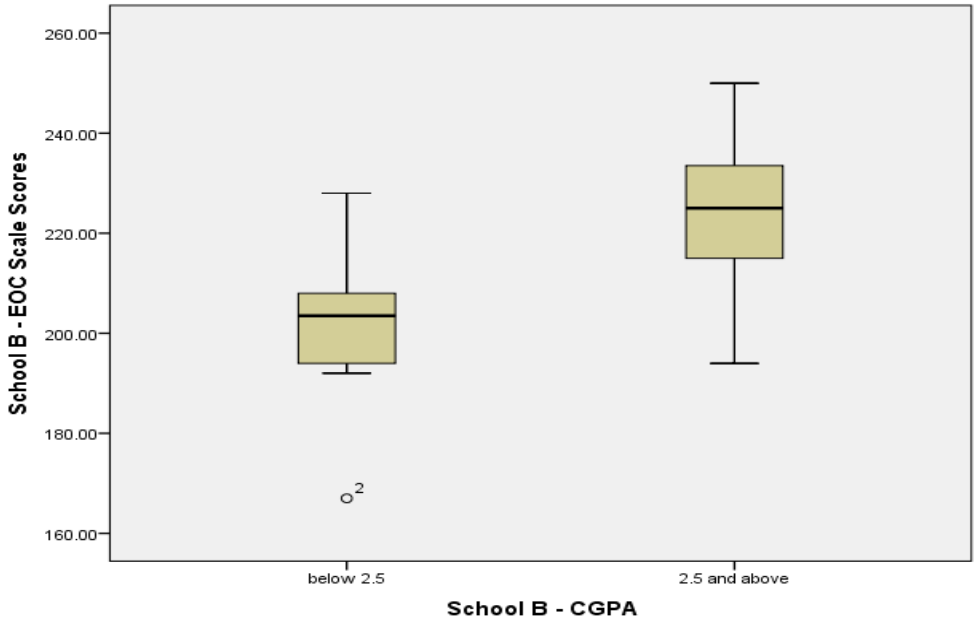
Pearson Product Moment Correlation Coefficient, PPT

		EOC scores-School B	*CGPA-School B*
EOC scores-School B	Pearson Correlation	1	.719**
	Sig. (2 tailed)		.000
	N	50	50
CGPA-School B	Pearson Correlation	.719**	1
	Sig. (2 tailed)	.000	
	N	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

The box plot in figure two visually displays the correlation between CGPA and EOC scale scores for students taking the test by paper-pencil method. Students below 2.5 CGPA have the lower EOC scale scores, while students above 2.5 CGPA have the higher scale scores. The mean for the category of CGPA 2.5 and above exceeds that of the mean for the below 2.5 CGPAs.

Figure 2 – Box Plot, School District B, PPT, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 3: For students taking the United States Government EOC examination by computer-based mode of delivery, there is no relationship between student cumulative GPA and United States Government EOC exam score.

According to Table 12, the Pearson Product Moment Correlation Coefficient demonstrates a strong correlation that is statistically significant [$r(48) = 0.752$, alpha = .05, $p < 0.0005$], therefore the researcher rejected the null hypothesis and data supports the alternative, indicating there is a strong positive relationship between students CGPA and EOC scores when taking an EOC exam using the CBT method.

Table 12

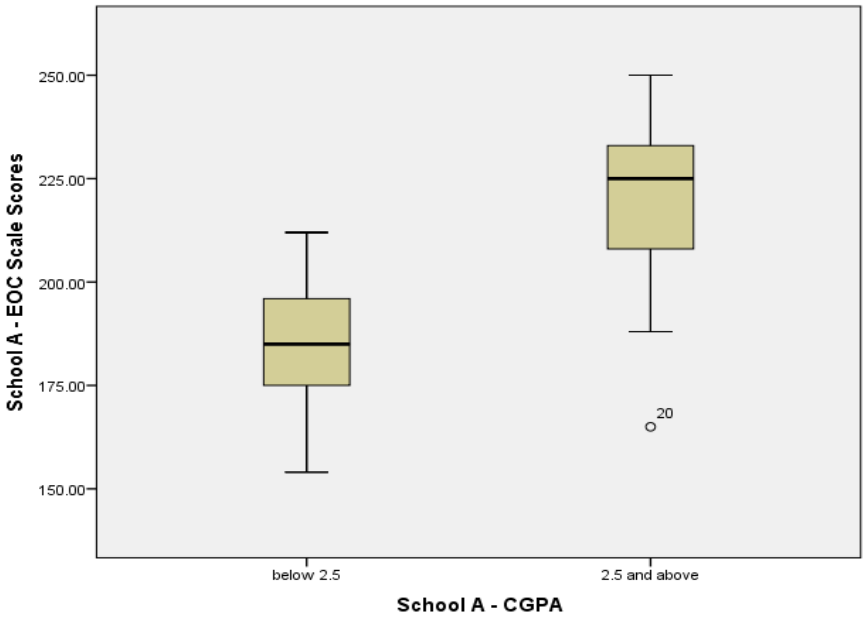
Pearson Product Moment Correlation Coefficient, CBT

		CGPA	*EOC*
CGPA	Pearson Correlation	1	.752**
	Sig. (2-tailed)		.000
	N	50	50
EOC	Pearson Correlation	.752**	1
	Sig. (2-tailed)	.000	
	N	50	50

** Correlation is significant at the 0.01 level (2-tailed). Alpha = 0.05

Figure 3 demonstrates an observable difference in means and the ranges of scores between the categories of CPGA below 2.5, and 2.5 and above.

Figure 3 – Box Plot, School District A, CBT, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 4: For students taking the United States Government EOC examination by paper-pencil mode of delivery, students with a cumulative GPA of 2.5 or higher will not score higher than students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 13, there is an observable difference between the means for the two samples. A Chi-Square test for Homogeneity was applied to data [X (50) = 0 and X-critical = 3.845] which allowed the researcher to not reject the null hypothesis. A Confidence Interval test for difference in means applied to data from the PPT administration of the EOC resulted in the range [-166.5, 123.9]. Since the test value of 0 is in the range, the researcher did not reject the null hypothesis. Data from this study indicates that there is no significant difference between students with CGPAs of 2.5 or

above and those with a CGPA of less than 2.5 on PPT scores on Government EOC exams.

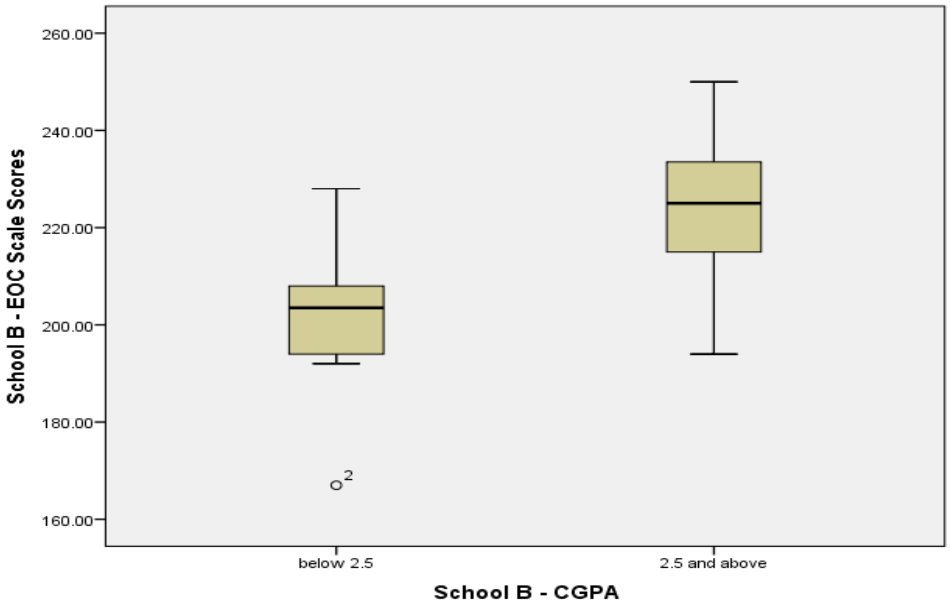
Table 13

Statistics for School District B, PPT, Null Hypothesis 4

	GPA	N	Mean	Std. Deviation	Std. Error Mean
School District B Scores	below 2.5	14	202.5000	14.93962	3.99278
	2.5 and above	36	223.8333	13.58676	2.26446

Figure 4 indicates there is a difference in means and an observable difference in low and high range of EOC scores based on CGPA.

Figure 4 – Box Plot, School District B, PPT, Relationship of CGPA and EOC Scale Scores



Null Hypothesis 5: For students taking the United States Government EOC examination by computer-based mode of delivery, students with a cumulative GPA below 2.5 will not score higher than students with a cumulative GPA of 2.5 or above, on the United States Government EOC exam.

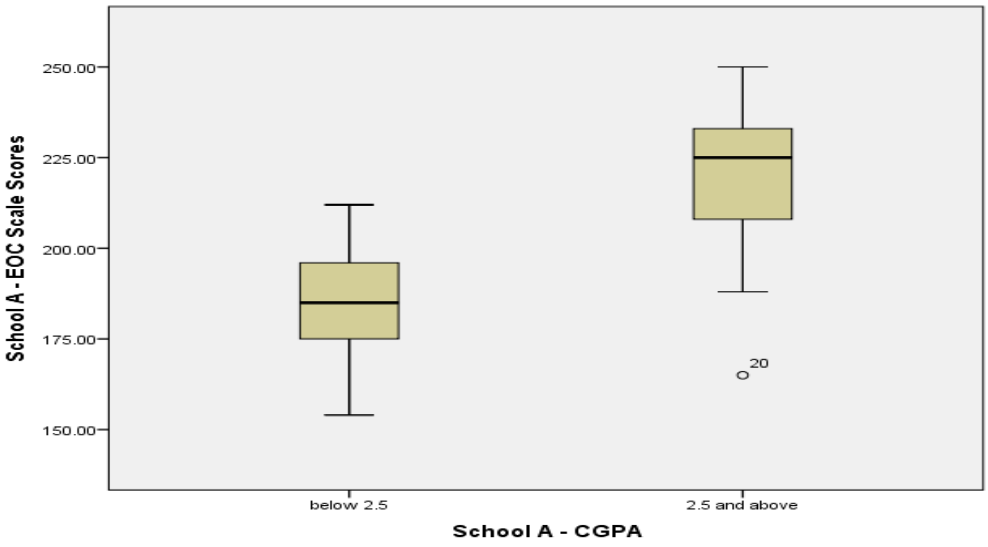
According to Table 14, there is an observable difference between the means for the two samples. A Chi-Square test for Homogeneity was applied to data [X (50) = 0 and X-critical = 3.845] which allowed the researcher to not reject the null hypothesis. A Confidence Interval test for difference in means applied to data from the CBT administration of the EOC resulted in the range [-260.3, 186.9]. Since the test value of 0 is in the range, the researcher did not reject the null hypothesis. Data from this study indicates that there is no significant difference between students with CGPAs of 2.5 or above and those with a CGPA of less than 2.5 on CBT scores on Government EOC exams.

Table 14

Statistics for School District A, CBT, Null Hypothesis 5

GPA-group	N	Mean	Std. Deviation	Std. Error Mean
EOC below 2.5	12	184.0833	16.41761	4.73936
EOC 2.5 and above	38	220.7368	20.12083	3.26403

Figure 5 - Box Plot, School Disrict A, CBT, Relationship of CGPA and EOC Scale Scores



Null hypotheses one, four and five were not rejected. Null hypotheses two and three were rejected and data supports the alternative hypotheses that there is a relationship between student CGPA and United States Government EOC exam scores whether taken PPT or CBT.

Data by Subgroups

Statistical Analysis, Gender, Female

Null Hypothesis 1a: For female students taking the United States Government EOC examination by paper-pencil mode of delivery, compared to female students taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam. A *t*-test for difference in means was conducted using female scale scores from both schools A (computer-based testing) and B (paper-pencil-based testing). According to the descriptive statistics in Table 15 the EOC scale scores of females at school B are significantly greater than the EOC scale scores of females at school A, which was supported by the *t*-test [t (91.33) = 2.356, p = 0.021]: therefore, the researcher rejected the null hypothesis.

Table 15

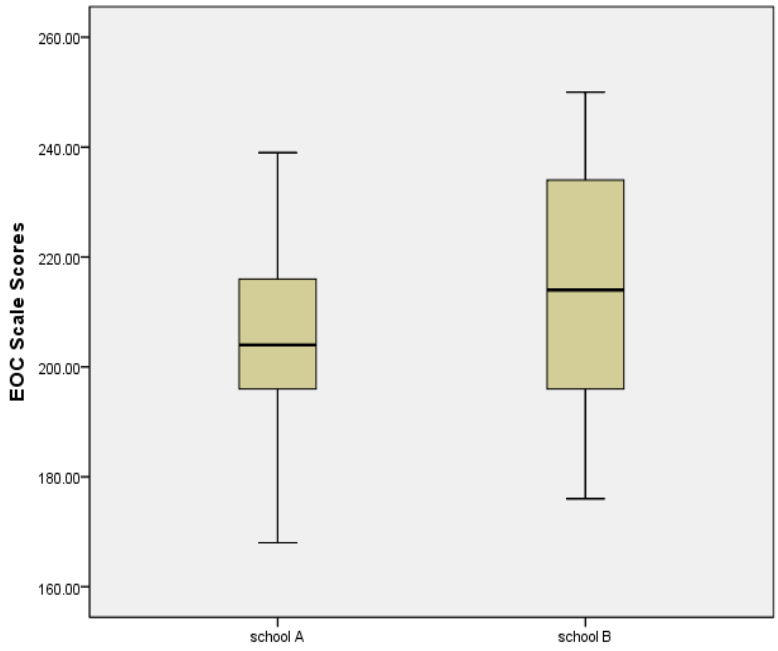
Descriptive Statistics Comparing Females, Schools A and B

School A & B	N	Mean	Std. Deviation	Std. Error Mean
EOC school A	50	204.8000	16.60710	2.34860
EOC school B	50	213.9600	21.90887	3.09838

Note: School A used computer-based and School B used paper-pencil-based testing.

The box plot in Figure 6 demonstrates the means and the range for School A (CBT) show higher EOC scale scores for School B.

Figure 6 – Box Plot, Comparison of Female EOC Scale Scores by School District
School District A = CBT School District B = PPT



Null Hypothesis 2a: For female students taking the United States Government EOC examination by paper-pencil mode of delivery, there is no relationship between female student cumulative GPA and United States Government EOC exam score.

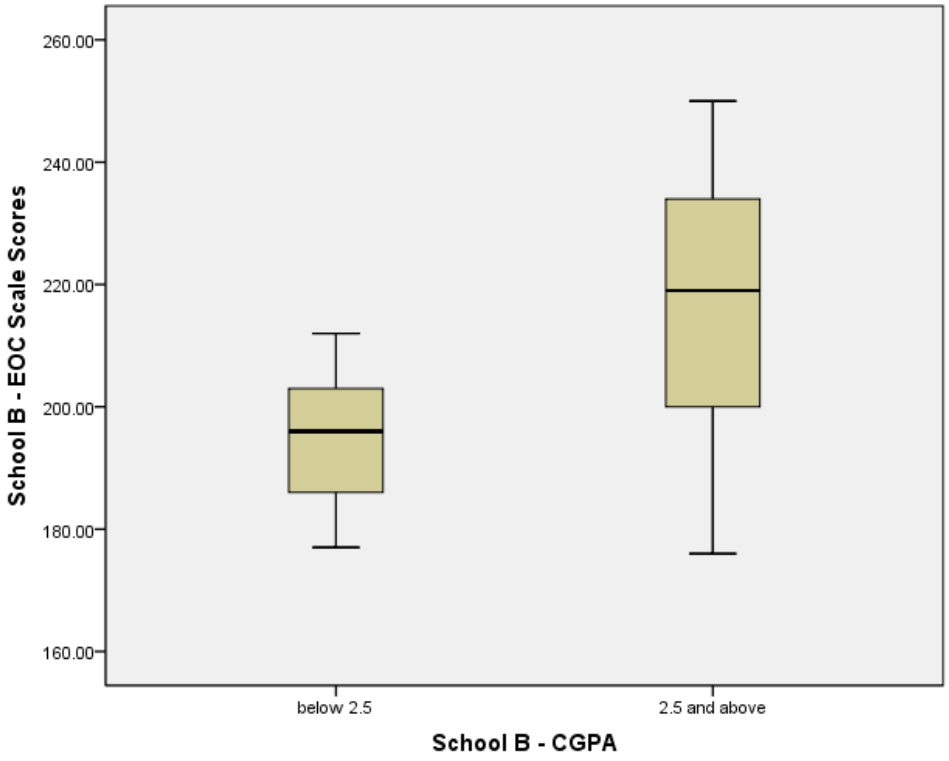
According to Table 16, the Pearson Product Moment Correlation demonstrates a strong correlation that is statistically significant [$R(48)=0.682$, $p<0.0005$] therefore the researcher rejected the null hypothesis and data supports the alternative, indicating there is a strong positive relationship between female students CGPA and EOC scores for female students taking the PPT form of the U.S. Government EOC.

Table 16

Pearson Product Moment Correlation Coefficient

		EOC scores-School B	CGPA-School B
EOC scores-School B	Pearson Correlation	1	.682**
	Sig. (2-tailed)		.000
	N	50	50
CGPA-School B	Pearson Correlation	.682**	1
	Sig. (2-tailed)	.000	
	N	50	50

Figure 7 – Box Plot, School District B, PPT, Females, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 3a: For female students taking the United States Government EOC examination by computer-based mode of delivery, there is no relationship between female student cumulative GPA and United States Government EOC exam score.

According to Table 17, the Pearson Product Moment Correlation Coefficient demonstrates a moderately strong correlation that is statistically significant [$R(48) = 0.568$, $p < 0.0005$] therefore the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between female students CGPA and EOC scores for female students taking a CBT.

Table 17

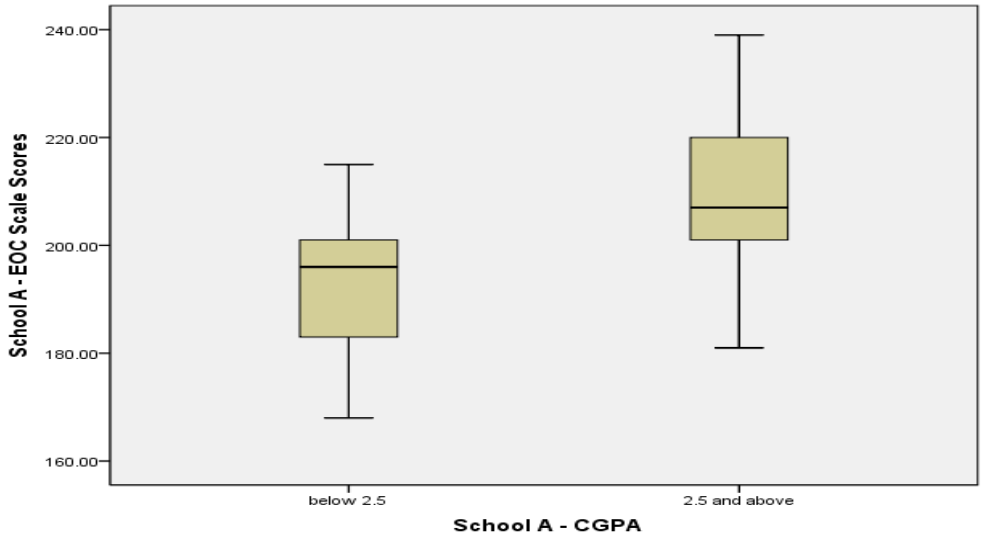
Pearson Product Moment Correlation Coefficient

		EOCscores-School A	CGPA-School A
EOC scores-School A	Pearson Correlation	1	.568**
	Sig. (2-tailed)		.000
	N	50	50
CGPA-School A	Pearson Correlation	.568**	1
	Sig. (2-tailed)	.000	
	N	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 8 demonstrates an observable difference in means and the range of scores between 2.5, and 2.5 and above CGPAs.

Figure 8 – Box Plot, School District A, CBT, Females, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 4a: For female students taking the United States Government EOC examination by paper-pencil mode of delivery, female students with a cumulative GPA of 2.5 or higher will not score higher than female students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 18, there is an observable difference between the means for the two samples. A Chi-Square test for Homogeneity was applied to data [X (50) = 0 and X-critical = 3.845] which allowed the researcher to not reject the null hypothesis. A Confidence Interval test for difference in means applied to data from the PPT administration of the EOC resulted in the range [-189.9, 235.3]. Since the test value of 0 is in the range, the researcher did not reject the null hypothesis. Data from this study indicates that there is no significant difference between female students with CGPAs of 2.5 or above and those with a CGPA of less than 2.5 on PPT scores on Government EOC exams.

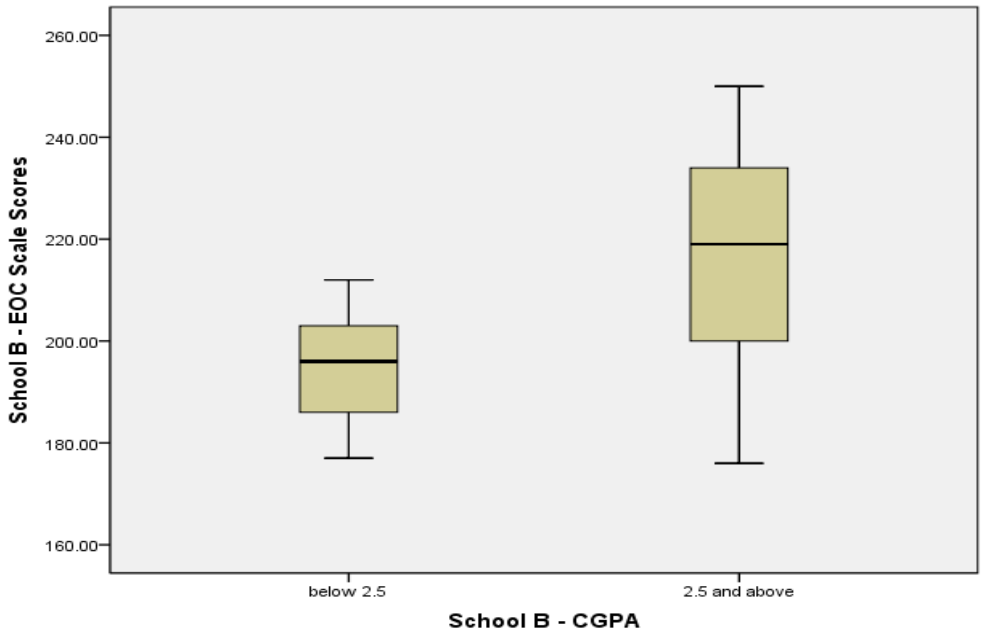
Table 18

Statistics for School District B, PPT, Females, Null Hypothesis 4

	GPA	N	Mean	Std. Deviation	Std. Error Mean
School B EOC Scores	2.5 and above	42	217.5952	21.54978	3.32520
	below 2.5	8	194.8750	12.01710	4.24869

Figure 9 indicates there is a difference in means and a noticeable difference in the high range of EOC scores based on CGPA.

Figure 9 – Box Plot, School District B, PPT, Females, Relationship of CGPA and EOC Scale Scores



Null Hypothesis 5a: For female students taking the United States Government EOC examination by computer-based mode of delivery, female students with a cumulative GPA of 2.5 or higher will not score higher than female students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

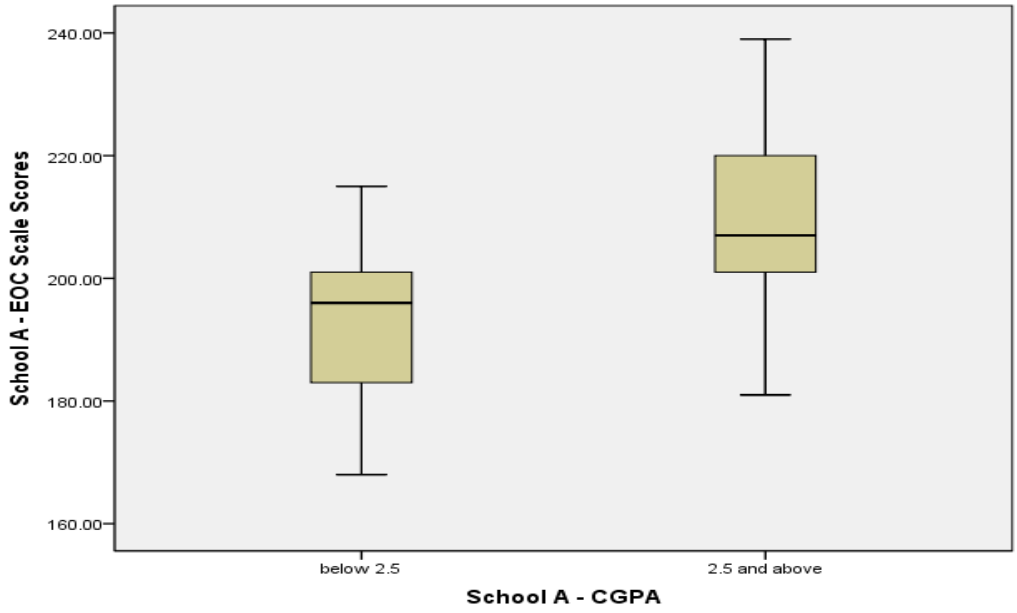
According to Table 19, there is an observable difference between the means for the two samples. A Chi-Square test for Homogeneity was applied to data [X (50) = 0 and X-critical = 3.845] which allowed the researcher to not reject the null hypothesis. A Confidence Interval test for difference in means applied to data from the CBT administration of the EOC resulted in the range [0-245.4, 279.6]. Since the test value of 0 is in the range, the researcher did not reject the null hypothesis. Data from this study indicates that there is no significant difference between female students with CGPAs of 2.5 or above and those with a CGPA of less than 2.5 on CBT scores on Government EOC exams.

Table 19

Statistics for School District A, CBT, Females, Null Hypothesis 5

CGPA group	N	Mean	Std. Deviation	Std. Error Mean
EOC 2.5 and above	37	209.2973	14.79652	2.43253
EOC below 2.5	13	192.0000	15.14926	4.20165

Figure 10 – Box Plot, School District A, CBT, Females, Relationship of CGPA and EOC Scale Scores



Statistical Analysis, Gender, Males

Null Hypothesis 1b: For male students taking the United States Government EOC examination by paper-pencil mode of delivery, compared to male students taking the examination by computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam.

According to the descriptive statistics in Table 20, the EOC scale scores of males at schools A and B shows a small difference between the EOC scale scores of males at school A and those of males at school B, which was supported by the t-test [$t(98) =$

1.191, $p = 0.236$]. The researcher did not reject the null hypothesis that there were no differences in student test scores based on mode of test administration.

Table 20

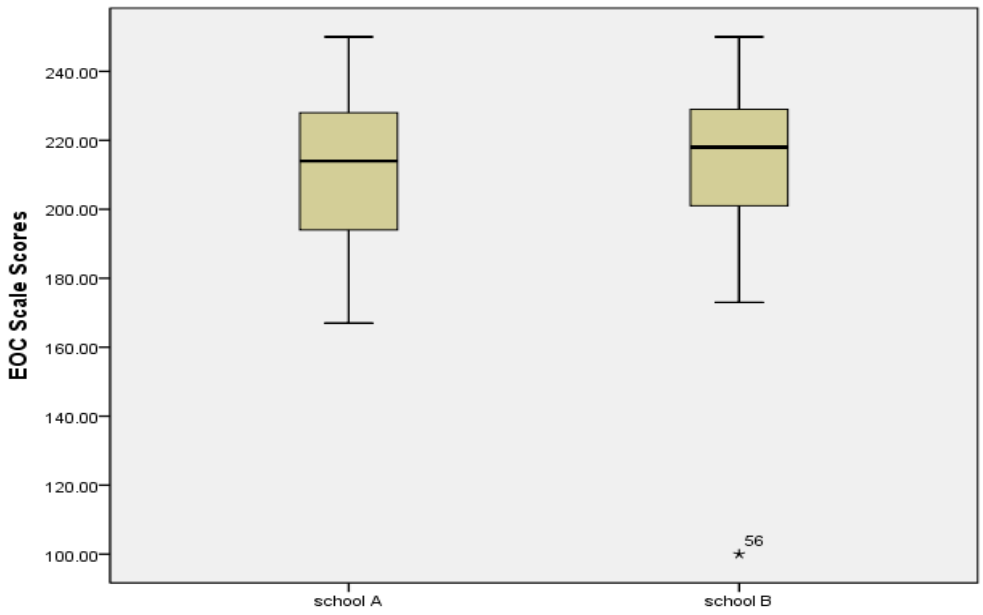
Descriptive Statistics Comparing Males, Schools A and B

School A & B	N	Mean	Std. Deviation	Std. Error Mean
EOC school A	50	210.4400	23.44264	3.31529
EOC school B	50	216.1600	24.56674	3.47426

Note: School A used computer-based and School B used paper-pencil-based testing.

The box plot in Figure 11 demonstrates there is very little difference in means and ranges of EOC Scale Scores for Schools A and B.

Figure 11 – Box Plot, Comparison of Male EOC Scale Scores by School District
School District A = CBT School District B = PPT



Null Hypothesis 2b: For male students taking the United States Government EOC examination by paper-pencil mode of delivery, there is no relationship between male student cumulative GPA and United States Government EOC exam score.

According to Table 21, the Pearson Product Moment Correlation Coefficient

demonstrates a strong correlation that is statistically significant [R(48)=0.615, p<0.0005] therefore the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between male students CGPA and EOC scores when taking the PPT form of the U.S. Government EOC.

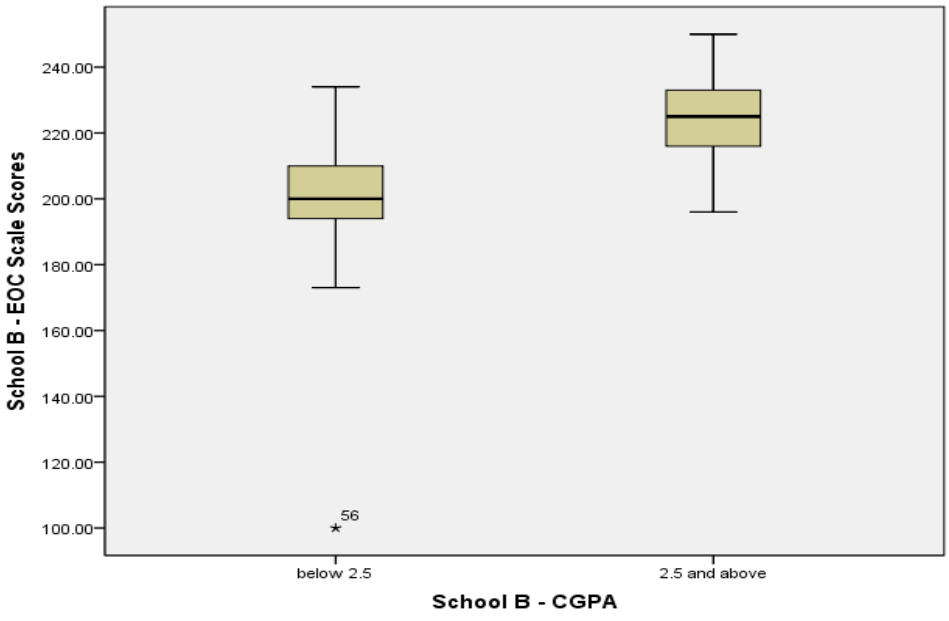
Table 21

Pearson Product Moment Correlation Coefficient

		CGPA-School B	EOC scores-School B
CGPA-School B	Pearson Correlation	1	.615**
	Sig. (2-tailed)		.000
	N	50	50
EOC scores-School B	Pearson Correlation	.615**	1
	Sig. (2-tailed)	.000	
	N	50	50

**.Correlation is significant at the 0.01 level (2-tailed).

Figure 12 – Box Plot, School District B, PPT, Males, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 3b: For male students taking the United States Government EOC examination by computer-based mode of delivery, there is no relationship between male

student cumulative GPA and United States Government EOC exam score.

According to Table 22, the Pearson Product Moment Correlation Coefficient demonstrates a strong relationship that is statistically significant [R(48) = 0.644, p <0.0005]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between male students CGPA and EOC scores when taking a CBT.

Table 22

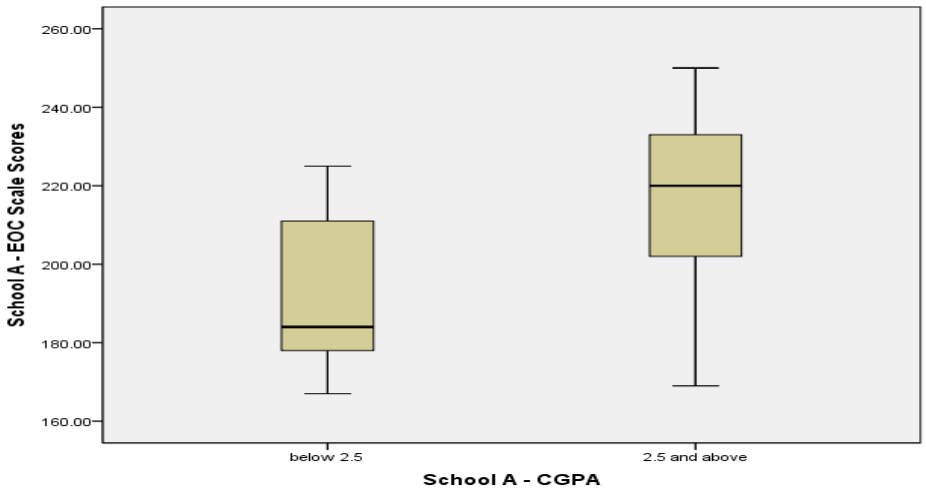
Pearson Product Moment Correlation Coefficient

		CGPA	EOC
CGPA	Pearson Correlation	1	.644**
	Sig. (2-tailed)		.000
	N	50	50
EOC	Pearson Correlation	.644**	1
	Sig. (2-tailed)	.000	
	N	50	50

**.Correlation is significant at the 0.01 level (2-tailed).

Figure 13 demonstrates a significant difference in means and range of scores between below 2.5 and 2.5 and above CGPAs.

Figure 13 – Box Plot, School District A, CBT, Males, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 4b: For male students taking the United States Government EOC examination by paper-pencil mode of delivery, male students with a cumulative GPA of 2.5 or higher will not score higher than male students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 23, the *t*-test for difference in means demonstrates the difference between the two groups is observable, which is supported by rejection of the null hypothesis [t (48) = 4.924, p = 0.0005]. Data supports the alternative, indicating the EOC scale scores of males at School B who had CGPAs of 2.5 and above are significantly greater than the EOC scale scores of males who had CGPAs below 2.5. Data from this study indicates that male students with CGPAs of 2.5 or higher will score higher than those with a CGPA of less than 2.5 on PPTs.

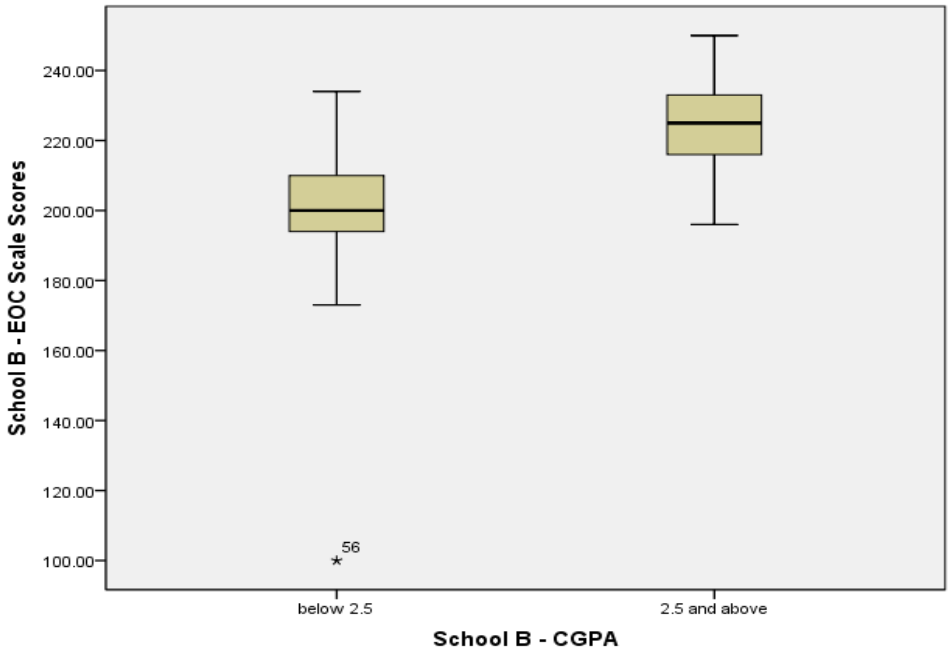
Table 23

Statistics for School District B, PPT, Males, Null Hypothesis 4

	GPA	N	Mean	Std. Deviation	Std. Error Mean
School District B Scores	2.5 and above	33	226.2727	14.11841	2.45770
	below 2.5	17	196.5294	28.79696	6.98429

Figure 14 indicates there is an observable difference in means and in the high range of EOC scores based on CGPA.

Figure 14 – Box Plot, School District B, PPT, Males, Relationship of CGPA and EOC Scale Scores



Null Hypothesis 5b: For male students taking the United States Government EOC examination by computer-based mode of delivery, male students with a cumulative GPA of 2.5 or higher will not score higher than male students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 24, the difference between the two groups is observable. The researcher rejected the null hypothesis and data supports the alternative [t (48) = 4.355, p <0.0005] indicating there is a contribution to a relationship between students CGPA and EOC scores when taking a CBT. The EOC scale scores of males at school A, who had CGPAs of 2.5 and above, were significantly greater than the EOC scale scores of males with CGPAs below 2.5.

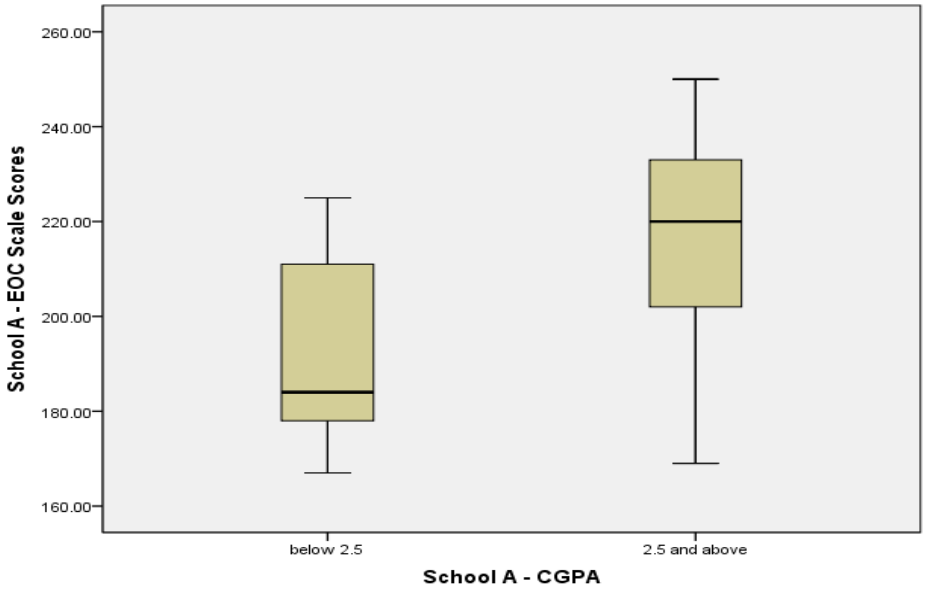
Table 24

Statistics for School District A, CBT, Males, Null Hypothesis 5

GPA	N	Mean	Std. Deviation	Std. Error Mean
Score 2.5 and above	34	218.9118	20.56677	3.52717
Score below 2.5	16	192.4375	18.87492	4.71873

According to Figure 15, the difference in means supports the alternative hypothesis that there is a relationship between students' CGPA and EOC scores, when taking a CBT form of the EOC exam.

Figure 15 – Box Plot, School District A, CBT, Males, Relationship of CGPA and EOC Scale Scores



Statistical Analysis, Eligible for FRL

Null Hypothesis 1c: For FRL eligible-students taking the United States Government EOC examination by paper-pencil mode of delivery, compared to FRL-eligible students taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam.

According to the descriptive statistics in Table 25 the EOC scale scores for those eligible to receive FRL at schools A and B show there is no statistical difference between the EOC scale scores for FRL at school A and those of FRL at school B, which was supported by the t-test [t (98) = 1.658, p=0.101]: therefore, the researcher did not reject the null hypothesis that there were no differences in student test scores based on mode of test administration.

Table 25

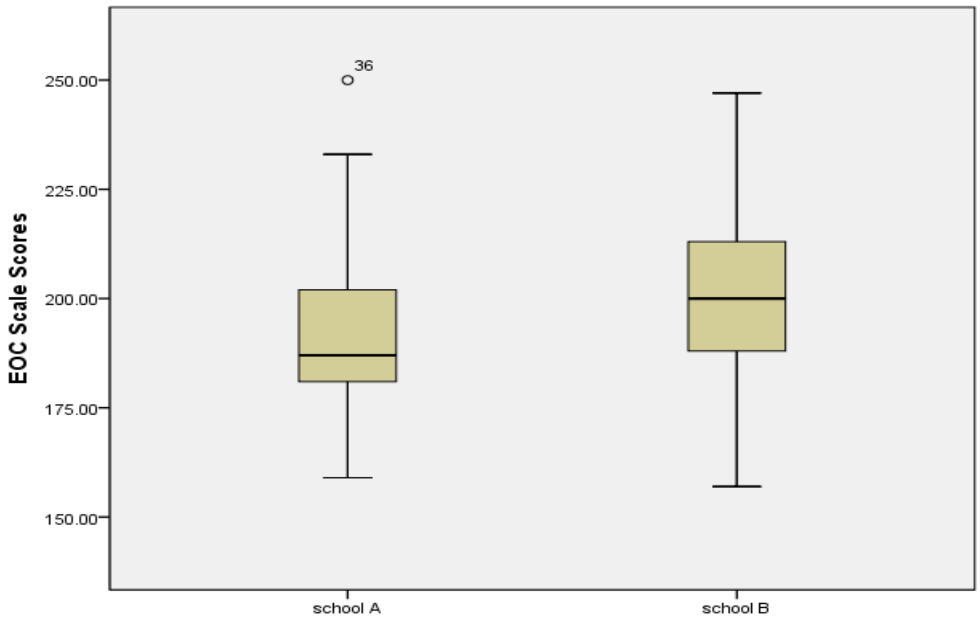
Descriptive Statistics Comparing Those Eligible for FRL, Schools A and B

	N	Mean	Std. Deviation	Std. Error Mean
EOC school A	50	192.4200	18.45247	2.60957
EOC school B	50	198.9200	20.69491	2.92670

Note: School A used computer-based and School B used paper-pencil-based testing.

The box plot in Figure 16 demonstrates there is very little difference in means and ranges of EOC Scale Scores for Schools A and B.

Figure 16 – Box Plot, Comparison of FRL EOC Scale Scores by School District
School District A = CBT School District B = PPT



Null Hypothesis 2c: For FRL eligible-students taking the United States Government EOC examination by paper-pencil mode of delivery, there is no relationship between FRL eligible student cumulative GPA and United States Government EOC exam score.

According to Table 26, the Pearson Product Moment Correlation Coefficient demonstrates a weak relationship that is statistically significant [$R(48) = 0.336$, $p<0.017$]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a weak positive relationship between FRL students' CGPA and EOC scores when taking the PPT form of the U.S. Government EOC.

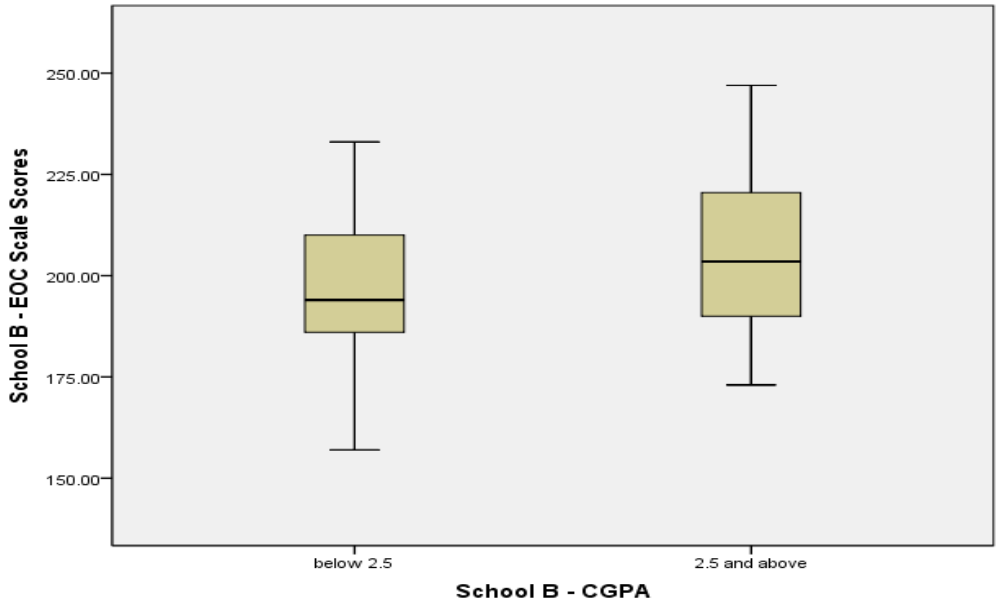
Table 26

Pearson Product Moment Correlation Coefficient

		EOC scores-School B	CGPA-School B
EOC scores-School B	Pearson Correlation	1	.336*
	Sig. (2-tailed)		.017
	N	50	50
CGPA-School B	Pearson Correlation	.336*	1
	Sig. (2-tailed)	.017	
	N	50	50

* Correlation is significant at the 0.05 level (2-tailed).

Figure 17 – Box Plot, School District B, PPT, FRL, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 3c: For FRL eligible-students taking the United States Government EOC examination by computer-based mode of delivery, there is no relationship between FRL eligible student cumulative GPA and United States Government EOC exam score.

According to Table 27 the Pearson Product Moment Correlation Coefficient demonstrates a moderately strong correlation that is statistically significant [$R(48) = 0.512$, $p < 0.0005$]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between FRL students' CGPA and EOC scores when taking a CBT type of EOC exam.

Table 27

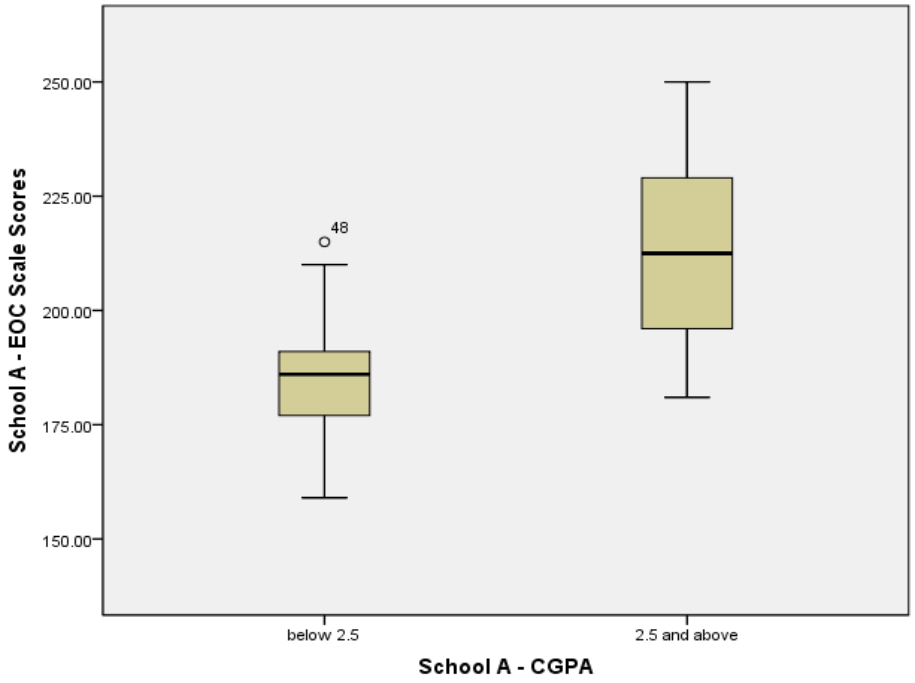
Pearson Product Moment Correlation Coefficient

		EOC	CGPA
CGPA	Pearson Correlation	1	.512**
	Sig. (2-tailed)		.000
	N	50	49
EOC	Pearson Correlation	.512**	1
	Sig. (2-tailed)	.000	
	N	49	49

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 18 demonstrates a significant difference in means and range of scores between below 2.5, and 2.5 and above CGPAs.

Figure 18 – Box Plot, School District A, CBT, FRL, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 4c: For FRL eligible-students taking the United States Government EOC examination by paper-pencil mode of delivery, FRL eligible students with a cumulative GPA of 2.5 or higher will not score higher than FRL eligible students

with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 28, the t-test for difference in means demonstrates there is an observable difference for FRL students between the two groups. Application of the *t*-test for difference in means [t (48) = 1.898, p = 0.06], allowed the researcher to not reject the null hypothesis.

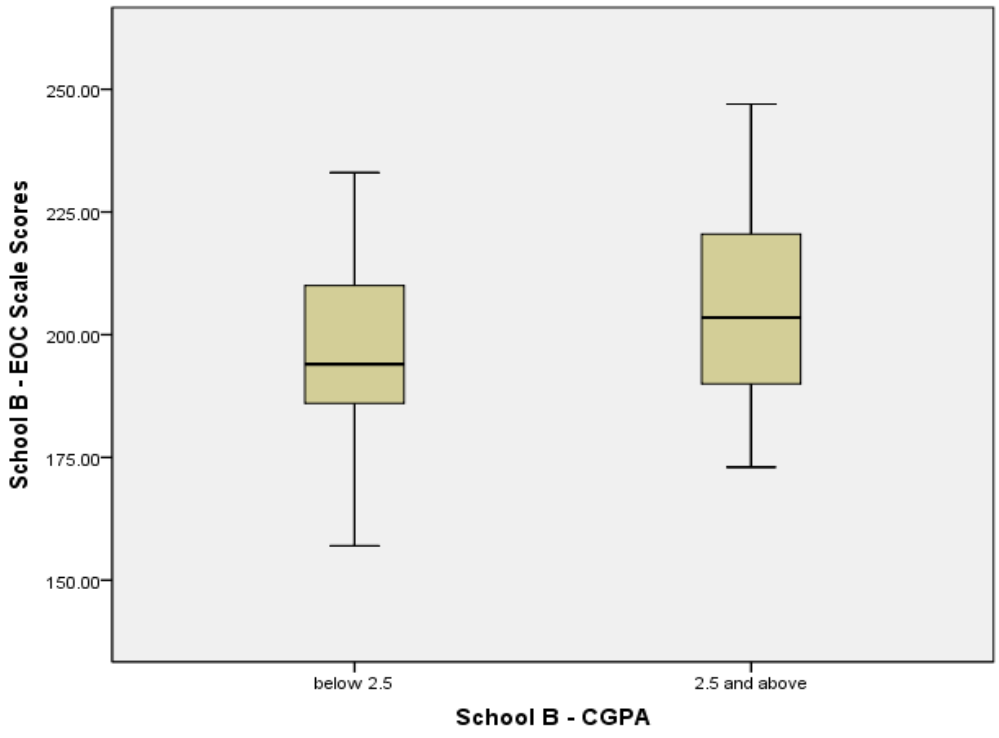
Table 28

Statistics for School District B, PPT, FRL, Null Hypothesis 4

GPA-group	N	Mean	Std. Deviation	Std. Error Mean
Score 2.5 and above	20	205.5500	21.30104	4.76306
Score below 2.5	30	194.5000	19.38672	3.53951

Figure 19 indicates there is an observable difference in means and range of EOC scores based on CGPA.

Figure 19 – Box Plot, School District B, PPT, FRL, Relationship of CGPA and EOC Scale Scores



Null Hypothesis 5c: For FRL eligible-students taking the United States Government EOC examination by computer-based mode of delivery, FRL eligible-students with a cumulative GPA of 2.5 or higher will not score higher than FRL eligible students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

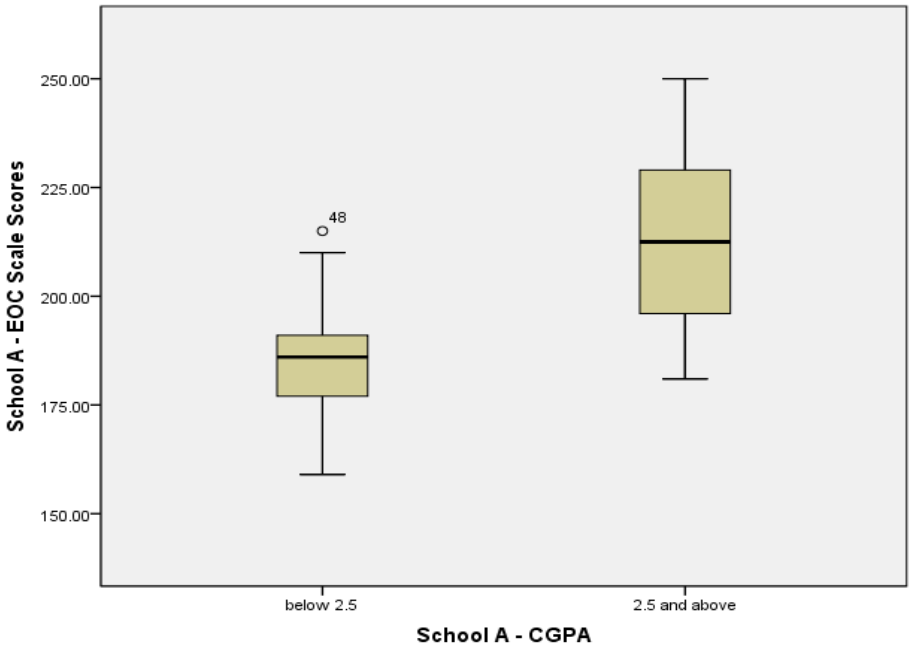
According to Table 29, there is an observable difference between the means for the two samples. A Chi-Square test for Homogeneity was applied to data [X (49) = 0 and X-critical = 3.845] which allowed the researcher to not reject the null hypothesis. A Confidence Interval test for difference in means was not applied to data from the PPT administration of the EOC due to insufficient sample size. A Confidence Interval test for difference in means applied to data from the CBT administration of the EOC resulted in the range [-275.9, 329.6]. Since the test value of 0 is in the range, the researcher did not reject the null hypothesis. Data from this study indicates that there is no significant difference between FRL students with CGPAs of 2.5 or above and those with a CGPA of less than 2.5 on CBT scores on Government EOC exams.

Table 29

Statistics for School District A, CBT, FRL, Null Hypothesis 5

GPA-group	N	Mean	Std. Deviation	Std. Error Mean
Score2.5 and above	12	212.8333	21.82923	6.30156
Score below 2.5	37	185.9730	11.62250	1.91073

Figure 20 – Box Plot, School District A, CBT, FRL, Relationship of CGPA and EOC Scale Scores



Statistical Analysis, Not Eligible for FRL

Null Hypothesis 1d: For students not FRL eligible taking the United States Government EOC examination by paper-pencil mode of delivery, compared to students not FRL eligible taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam.

According to the descriptive statistics in Table 30, the EOC scale scores for those not eligible to receive FRL at schools A and B show there is a small observable difference between the EOC scale scores for students not eligible for FRL at school A and those not eligible for FRL at school B. This observation was supported statistically with results from application of a t-test for difference in means [t (98) = 1.157, p = 0.250]. The researcher did not reject the null hypothesis that there were no differences in student test scores based on mode of test administration.

Table 30

Descriptive Statistics Comparing Those Not Eligible for FRL, Schools A and B

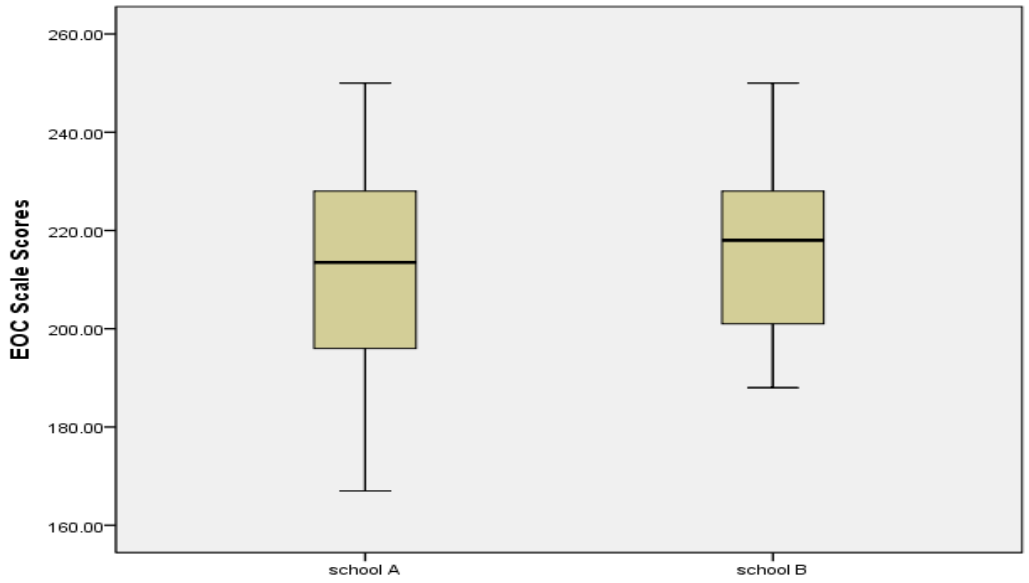
	N	Mean	Std. Deviation	Std. Error Mean
EOC school A	50	213.1000	22.43016	3.17210
EOC school B	50	217.8200	18.14195	2.56566

Note: School A used computer-based and School B used paper-pencil-based testing.

The box plot in Figure 21 demonstrates there is very little difference in means and range of EOC Scale Scores for those not eligible for FRL at Schools A and B

Figure 21 – Box Plot, Comparison of Not Eligible FRL EOC Scale Scores by School District

School District A = CBT School District B = PPT



Null Hypothesis 2d: For students not FRL-eligible, taking the United States Government EOC examination by paper-pencil mode of delivery, there is no relationship between not FRL-eligible student cumulative GPA and United States Government EOC exam score.

According to Table 31, the Pearson Product Moment Correlation Coefficient

demonstrates a moderate relationship that is statistically significant [R(48) = 0.552, p<0.0005]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between not eligible FRL students' CGPA and EOC scores when taking the PPT form of the U.S. Government EOC.

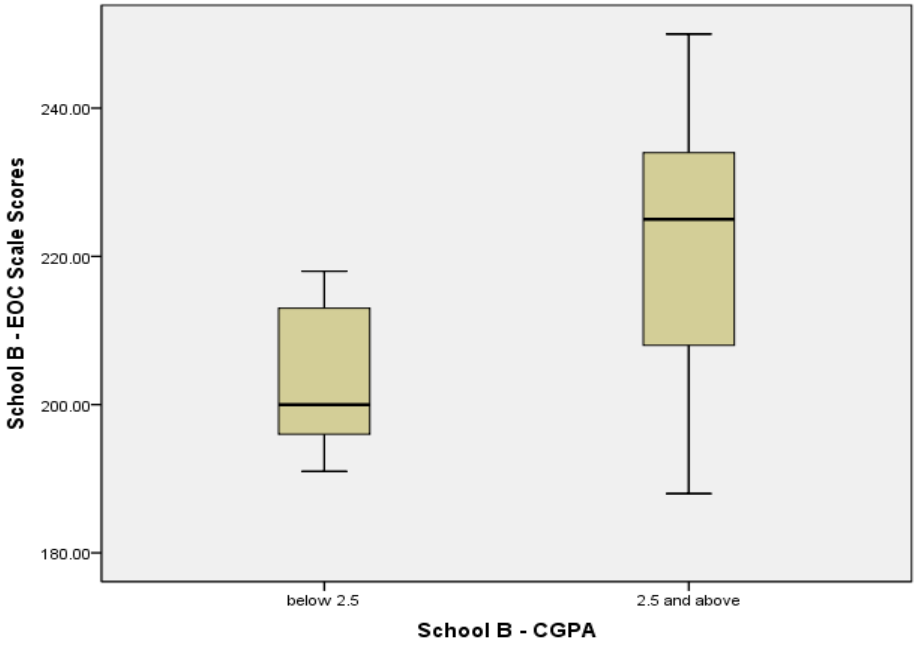
Table 31

Pearson Product Moment Correlation Coefficient

		CGPA-School B	EOC scores-School B
CGPA-School B	Pearson Correlation	1	.552**
	Sig. (2-tailed)		.000
	N	50	50
EOC scores-School B	Pearson Correlation	.552**	1
	Sig. (2-tailed)	.000	
	N	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 22 – Box Plot, School District B, PPT, Not Eligible for FRL, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 3d: For students not FRL-eligible, taking the United States Government EOC examination by computer-based mode of delivery, there is no

relationship between not FRL-eligible student cumulative GPA and United States Government EOC exam score.

According to Table 31, the Pearson Product Moment Correlation Coefficient demonstrates a moderately strong relationship that is statistically significant [$R(48) = 0.540$, $p < 0.0005$]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between not eligible for FRL students' CGPA and EOC scores when taking a CBT type EOC exam.

Table 32

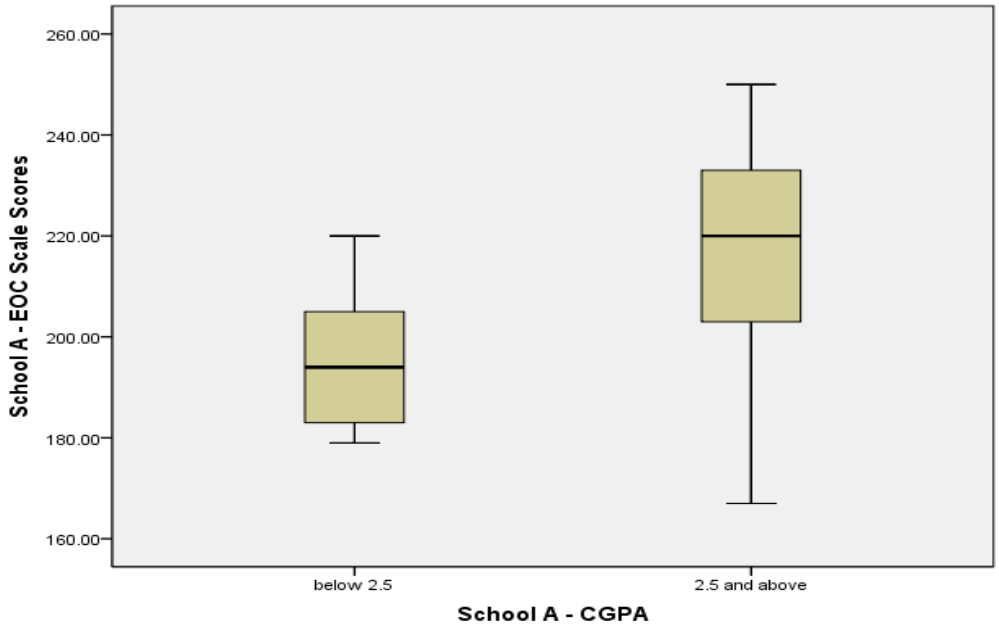
Pearson Product Moment Correlation Coefficient

		CGPA	EOC
CGPA	Pearson Correlation	1	.540**
	Sig. (2-tailed)		.000
	N	50	50
EOC	Pearson Correlation	.540**	1
	Sig. (2-tailed)	.000	
	N	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 23 demonstrates a significant difference in means and range of scores between below 2.5, and 2.5 and above CGPAs.

Figure 23 – Box Plot, School District A, CBT, Not Eligible for FRL, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 4d: For students not FRL eligible, taking the United States Government EOC examination by paper-pencil mode of delivery, students not FRL eligible, with a cumulative GPA of 2.5 or higher will not score higher than students not FRL eligible with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 33, the *t*-test for difference in means demonstrates there is an observable difference for not eligible FRL students between the two CGPA groups. Application of a *t*-test for difference in means [t (48) = 2.778, p = 0.008] allowed the researcher to not reject the Null Hypothesis. EOC scale scores of students not eligible to receive FRL who had CGPAs of 2.5 and above are significantly greater than the EOC scale scores of those who had CGPAs below 2.5. These results were verified with a Chi Square Test for Homogeneity [X=0; X-critical = 3.845] and Confidence Interval Tests for Difference in Means [-122.7, 157.4].

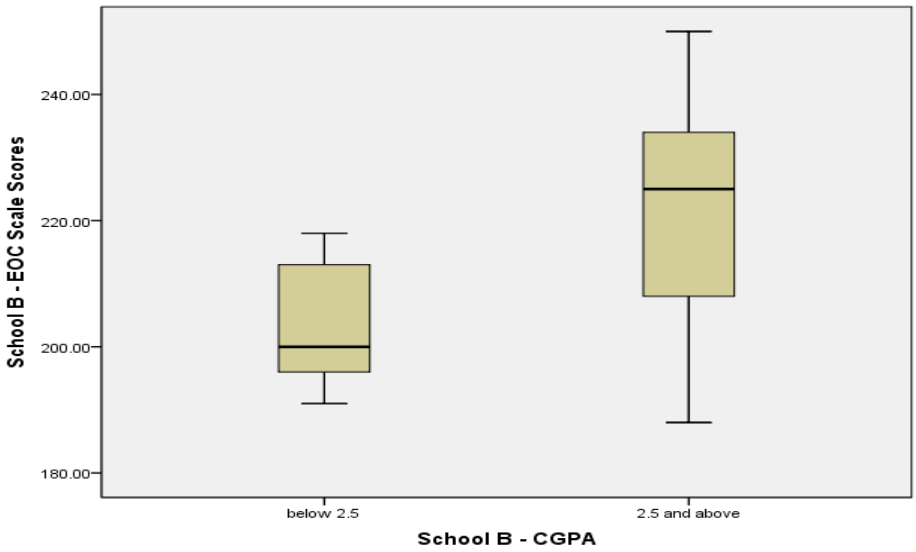
Table 33

Statistics for School District B, PPT, Not Eligible for FRL, Null Hypothesis 4

	CGPA-group	N	Mean	Std. Deviation	Std. Error Mean
School District B Scores	2.5 and above	41	220.9512	18.11898	2.82971
	below 2.5	9	203.5556	9.76103	3.25368

Figure 24 indicates there is a very noticeable difference in means and ranges of EOC scores based on CGPA.

Figure 24 – Box Plot, School District B, PPT, Not Eligible for FRL, Relationship of CGPA and EOC Scale Scores



Null Hypothesis 5d: For students not FRL eligible, taking the United States Government EOC examination by computer-based mode of delivery, students not FRL eligible with a cumulative GPA of 2.5 or higher, will not score higher than students not FRL eligible with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 34, the difference between the two CGPA groups is observable. The researcher rejected the null hypothesis and data supports the alternative [$t(48) = 3.051$, p

<0.004] indicating there is a contribution to a relationship between not eligible FRL students CGPA and EOC scores when taking a CBT. The EOC scale scores for not eligible for FRL at school A, who had CGPAs of 2.5 and above, are significantly greater than the EOC scale scores of not eligible FRL with CGPAs below 2.5. These results were verified with a Chi Square Test for Homogeneity [X=0; X-critical = 3.845] and Confidence Interval Tests for Difference in Means [9.91, 34.6].

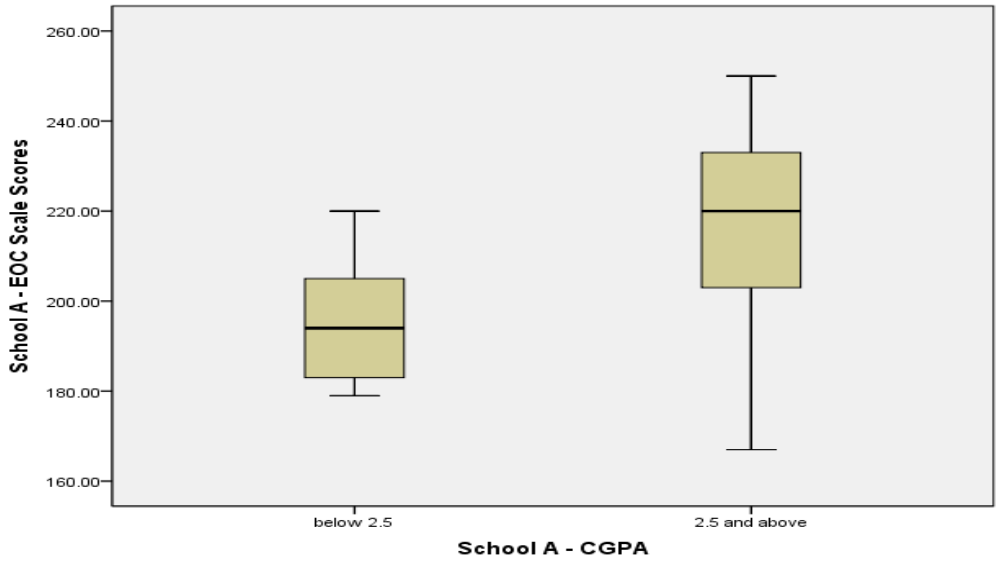
Table 34

Statistics for School District A, CBT, Not Eligible for FRL, Null Hypothesis 5

	CGPA-group	N	Mean	Std. Deviation	Std. Error Mean
EOC	2.5 and above	40	217.5750	22.09291	3.49320
	below 2.5	10	195.2000	13.38158	4.23163

According to Figure 25, the difference in means supports the alternative hypothesis that there is a relationship for not eligible FRL students, between students' CGPA and EOC scores, when taking a CBT form of the EOC exam.

Figure 25 – Box Plot, School District A, CBT, Not Eligible FRL, Relationship of CGPA and EOC Scale Scores



Statistical Analysis, With IEP

Null Hypothesis 1e: For IEP students taking the United States Government EOC examination by paper-pencil mode of delivery, compared to IEP students taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam.

According to the descriptive statistics in Table 35, the EOC scale scores for those with IEPs at schools A and B show there is a small observable difference between the EOC scale scores for IEP students at schools A and B. Application of a *t*-test for difference in means [t (98) = 1.124, p = 0.264] allowed the researcher to not reject the null hypothesis.

Table 35

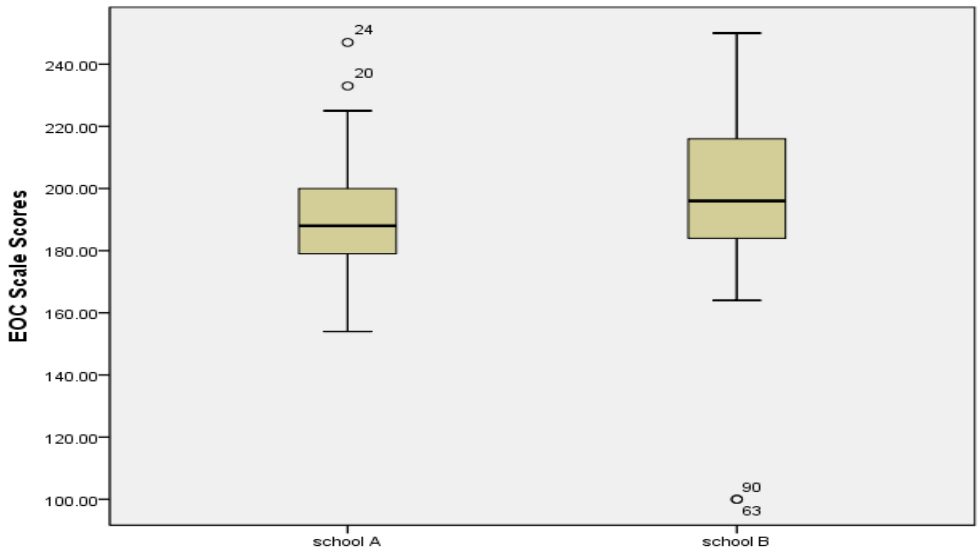
Descriptive Statistics Comparing Those With IEPs, Schools A and B

	N	Mean	Std. Deviation	Std. Error Mean
EOC school A	50	190.7200	18.65590	2.63834
EOC school B	50	196.1800	28.84136	4.07878

Note: School A used computer-based and School B used paper-pencil-based testing.

The box plot in Figure 26 demonstrates there is very little difference in means and ranges of EOC Scale Scores for IEP students at Schools A and B

Figure 26 – Box Plot, Comparison of IEP EOC Scale Scores by School District
School District A = CBT School District B = PPT



Null Hypothesis 2e: For IEP students taking the United States Government EOC examination by paper-pencil mode of delivery, there is no relationship between with an IEP student cumulative GPA and United States Government EOC exam score.

According to Table 36, the Pearson Product Moment Correlation Coefficient demonstrates a moderately strong relationship that is statistically significant [$R(48) = 0.553$, $p<0.0005$]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between IEP students' CGPA and EOC scores when taking the PPT form of the U.S. Government EOC.

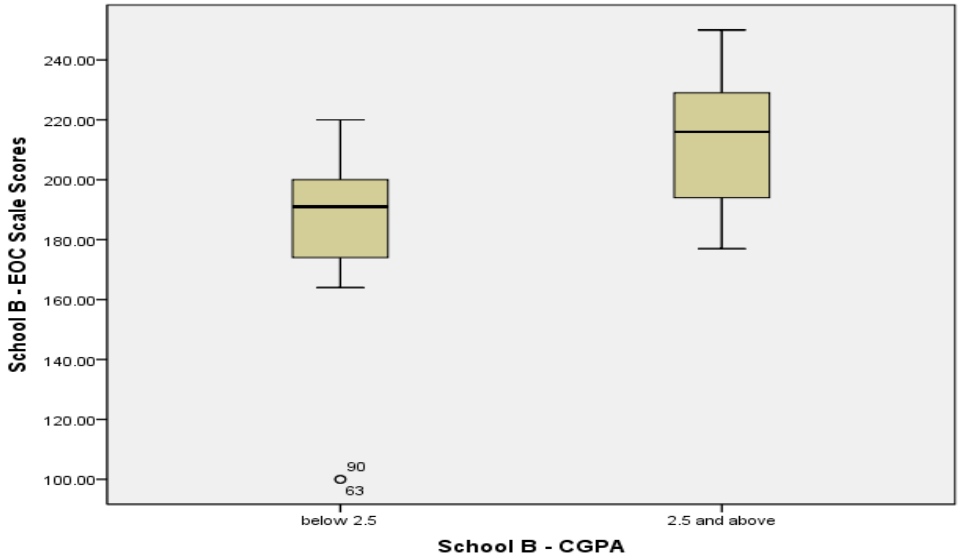
Table 36

Pearson Product Moment Correlation Coefficient

		EOC scores-School B	CGPA-School B
EOC scores-School B	Pearson Correlation	1	.553**
	Sig. (2-tailed)		.000
	N	50	50
CGPA-School B	Pearson Correlation	.553**	1
	Sig. (2-tailed)	.000	
	N	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 27 – Box Plot, School District B, PPT, IEP, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 3e: For IEP students taking the United States Government EOC examination by computer-based mode of delivery, there is no relationship between an IEP student cumulative GPA and United States Government EOC exam score.

According to Table 37, the Pearson Product Moment Correlation Coefficient demonstrates a mild relationship that is statistically significant [$R(48) = 0.461$, $p < 0.001$]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between IEP students' CGPA and EOC scores

when taking a CBT.

Table 37

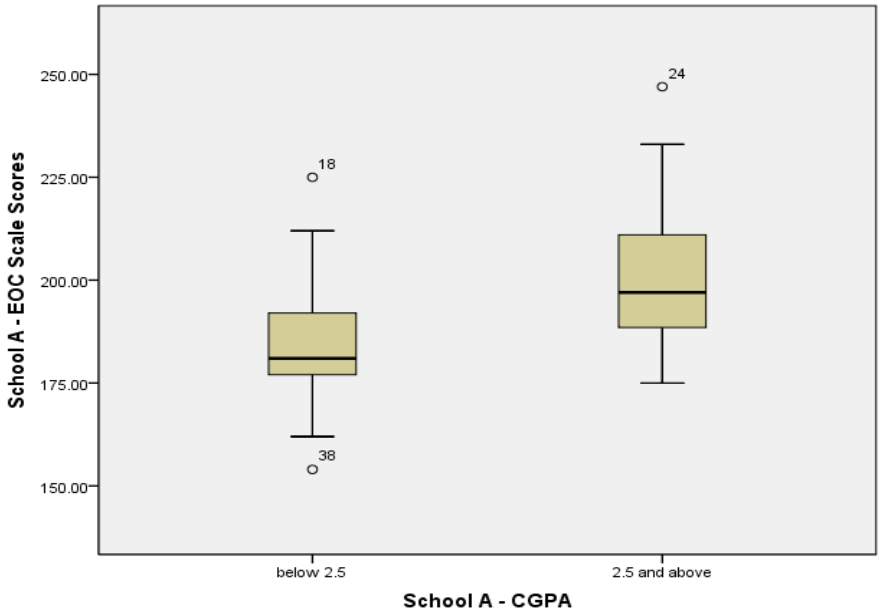
Pearson Product Moment Correlation Coefficient

		EOC	CGPA
EOC	Pearson Correlation	1	.461**
	Sig. (2-tailed)		.001
	N	50	49
CGPA	Pearson Correlation	.461**	1
	Sig. (2-tailed)	.001	
	N	49	49

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 28 demonstrates a difference in means and ranges of scores between below 2.5, and 2.5 and above CGPAs.

Figure 28 – Box Plot, School District A, CBT, IEP, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 4e: For IEP students taking the United States Government EOC examination by paper-pencil mode of delivery, IEP students with a cumulative GPA of 2.5 or higher will not score higher than IEP students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 38, the *t*-test for difference in means demonstrates there is an observable difference for IEP students between the two CGPA groups. Application of the t-test for difference in means [t (48) = 3.535, p = 0.001] allowed the researcher to reject the null hypothesis. EOC scale scores of IEP students who had CGPAs of 2.5 and above are significantly greater than the EOC scale scores of those who had CGPAs below 2.5

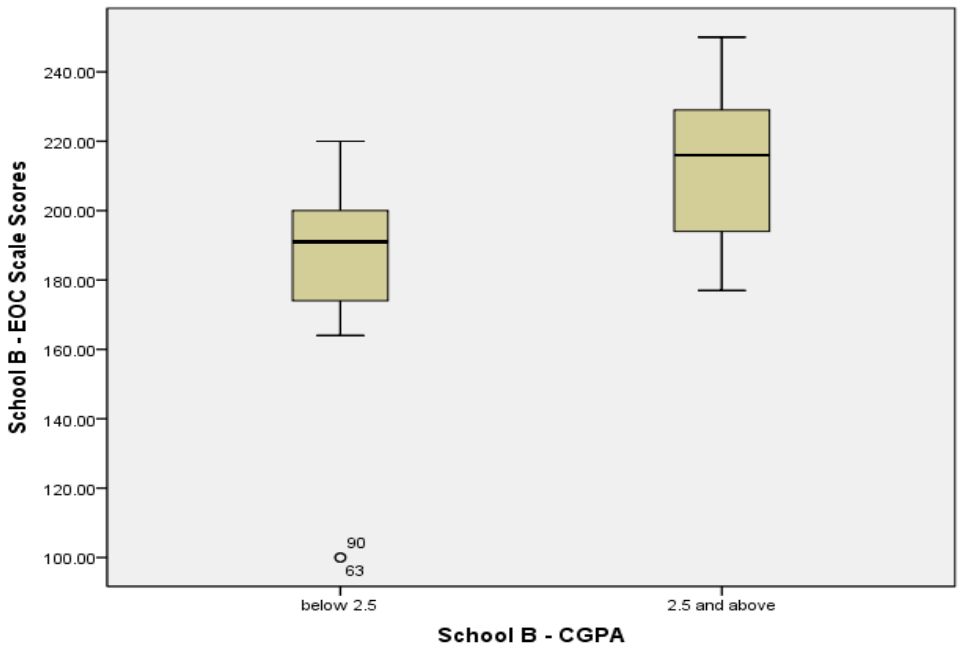
Table 38

Statistics for School District B, PPT, IEP, Null Hypothesis 4

	CGPA-group	N	Mean	Std. Deviation	Std. Error Mean
School District B Scores	2.5 and above	21	211.4286	22.42002	4.89245
	below 2.5	29	185.1379	28.21312	5.23905

Figure 29 indicates there is an observable difference in means and ranges of EOC scores based on CGPA.

Figure 29 – Box Plot, School District B, PPT, IEP, Relationship of CGPA and EOC Scale Scores



Null Hypothesis 5e: For IEP students taking the United States Government EOC examination by computer-based mode of delivery, IEP students with a cumulative GPA of 2.5 or higher will not score higher than IEP students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 39, the difference between the two CGPA groups is observable. The researcher rejected the null hypothesis and data supports the alternative [t (47) = 3.063, p = 0.004] indicating there is a contribution to a relationship between IEP students CGPA and EOC scores when taking a CBT. The EOC scale scores for IEP students at school A, who had CGPAs of 2.5 and above, are significantly greater than the EOC scale scores of IEP students with CGPAs below 2.5.

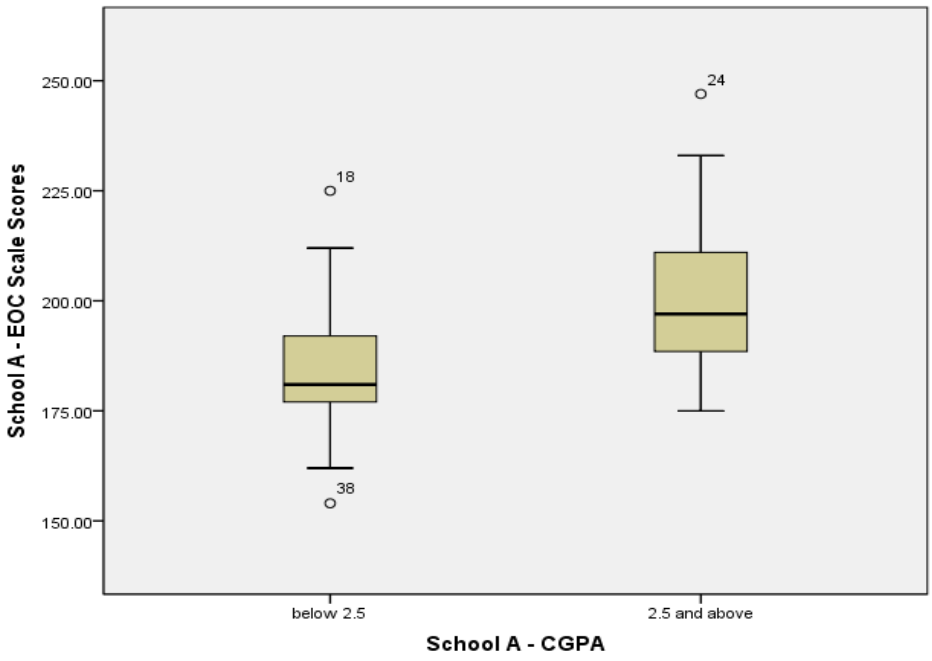
Table 39

Statistics for School District A, CBT, IEP, Null Hypothesis 5

	CGPA-group	N	Mean	Std. Deviation	Std. Error Mean
EOC	2.5 and above	20	200.2500	19.34690	4.32610
	below 2.5	29	185.2414	14.93953	2.77420

According to Figure 30, the difference in means supports the alternative hypothesis that there is a relationship for IEP students, between students' CGPA and EOC scores, when taking a CBT form of the EOC exam.

Figure 30 – Box Plot, School District A, CBT, IEP, Relationship of CGPA and EOC Scale Scores



Statistical Analysis, No IEP

Null Hypothesis 1f: For non-IEP students taking the United States Government EOC examination by paper-pencil mode of delivery, compared to non-IEP students taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam.

According to the descriptive statistics in Table 40, the EOC scale scores for those with no IEPs at schools A and B show there is a small observable difference between the EOC scale scores for students without IEPs at schools A and B. Application of the *t*-test for difference in means [t (92.94) = 1.706, p = 0.2091] allowed the researcher to not reject the null hypothesis.

Table 40

Descriptive Statistics Comparing Those Without IEPs, Schools A and B

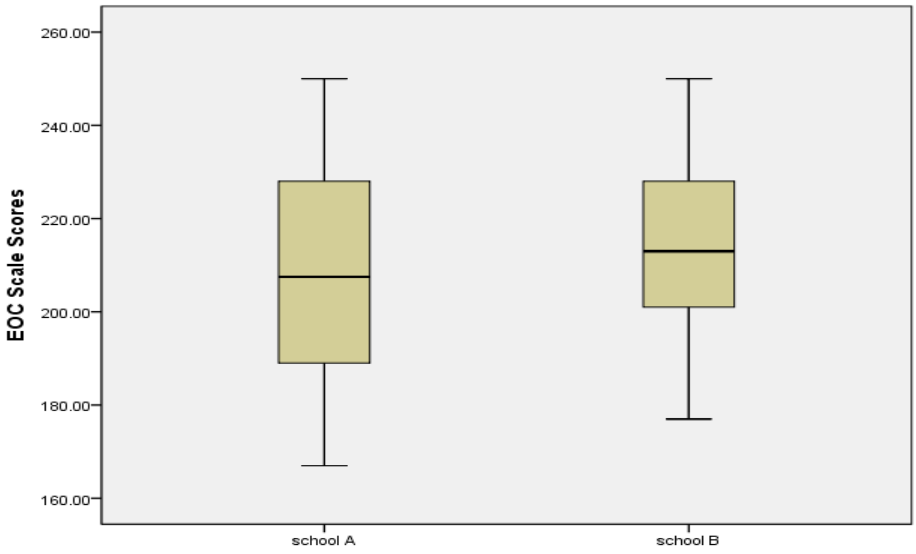
		N	Mean	Std. Deviation	Std. Error Mean
EOC	school A	50	208.5000	23.69384	3.35081
	school B	50	215.7800	18.68240	2.64209

Note: School A used computer-based and School B used paper-pencil-based testing.

The box plot in Figure 31 demonstrates there is very little difference in means and ranges of EOC Scale Scores for students without IEPs at Schools A and B

Figure 31 – Box Plot, Comparison No IEP EOC Scale Scores by School District

School District A = CBT School District B = PPT



Null Hypothesis 2f: For non-IEP students taking the United States Government EOC examination by paper-pencil mode of delivery, there is no relationship between with a non-IEP student cumulative GPA and United States Government EOC exam score.

According to Table 41, the Pearson Product Moment Correlation Coefficient demonstrates a strong relationship that is statistically significant [$R(48) = 0.664$, $p<0.0005$]: therefore, the researcher rejected the null hypothesis and data supports the

alternative, indicating there is a positive relationship between non-IEP students' CGPA and EOC scores when taking the PPT form of the U.S. Government EOC.

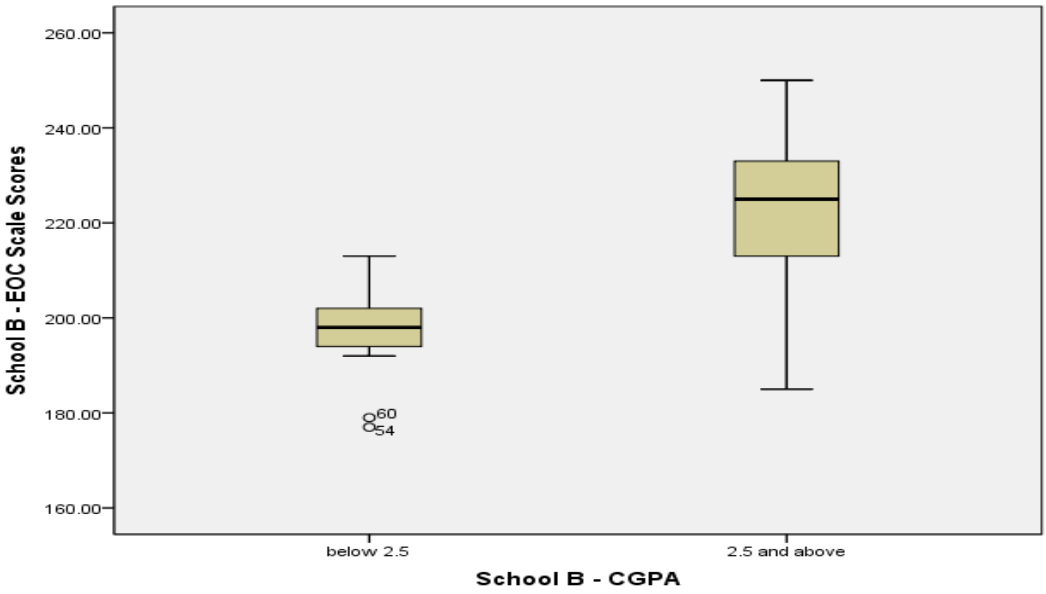
Table 41

Pearson Product Moment Correlation Coefficient

		EOC scores-School B	CGPA-School B
EOC scores-School B	Pearson Correlation	1	.664**
	Sig. (2-tailed)		.000
	N	50	50
CGPA-School B	Pearson Correlation	.664**	1
	Sig. (2-tailed)	.000	
	N	50	50

**.Correlation is significant at the 0.01 level (2-tailed).

Figure 32 – Box Plot, School District B, PPT, No IEP, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 3f: For non-IEP students taking the United States Government EOC examination by computer-based mode of delivery, there is no relationship between with a non-IEP student cumulative GPA and United States Government EOC exam score.

According to Table 42, the Pearson Product Moment Correlation Coefficient demonstrates a strong relationship that is statistically significant [$R(46) = 0.747$, p

<0.0005]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between non-IEP students' CGPA and EOC scores when taking a CBT.

Table 42

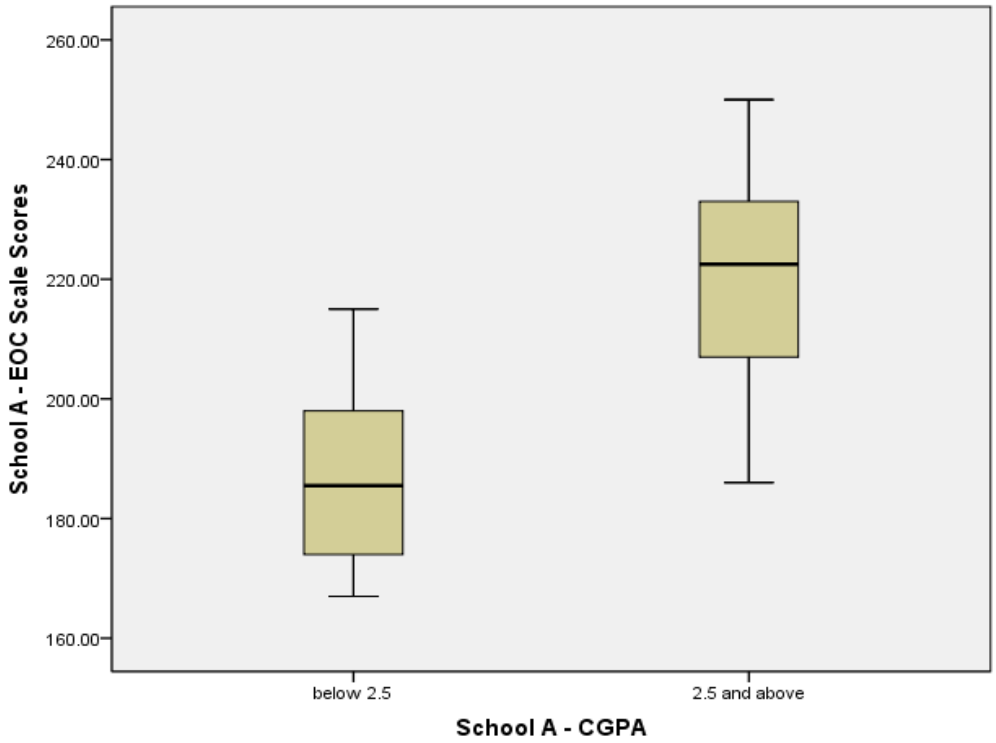
Pearson Product Moment Correlation Coefficient

		EOC	CGPA
EOC	Pearson Correlation	1	.747**
	Sig. (2-tailed)		.000
	N	50	48
CGPA	Pearson Correlation	.747**	1
	Sig. (2-tailed)	.000	
	N	48	48

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 33 demonstrates a significant difference in means and ranges of scores between below 2.5, and 2.5 and above CGPAs.

Figure 33 – Box Plot, School District A, CBT, No IEP, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 4f: For non-IEP students taking the United States Government EOC examination by paper-pencil mode of delivery, non-IEP students with a cumulative GPA of 2.5 or higher will not score higher than non-IEP students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 43, there is an observable difference between the means for the two samples. A Chi-Square test for Homogeneity applied to data from the PPT administration of the EOC resulted in [X (50) = 0 and X-critical = 3.845] which allowed the researcher to not reject the null hypothesis. A Confidence Interval test for difference in means resulted in the range [-88.5, 139.5]. Since the test value of 0 is in the range, the researcher did not reject the null hypothesis. Data from this study indicates that there is no significant difference between non-IEP students with CGPAs of 2.5 or above and those with a CGPA of less than 2.5 on PPT scores on Government EOC exams.

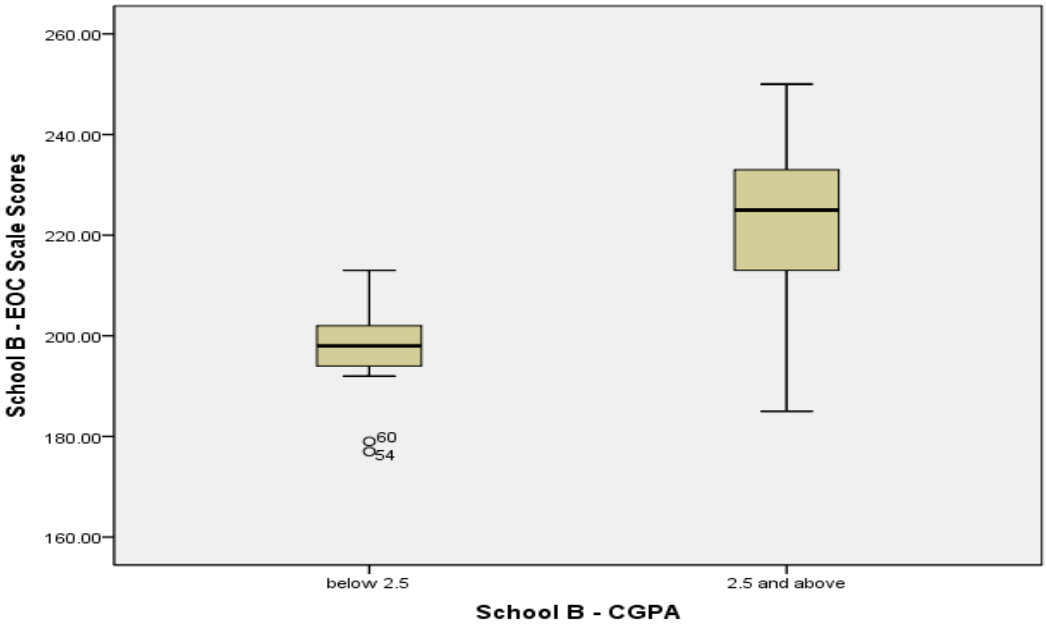
Table 43

Statistics for School District B, PPT, No IEP, Null Hypothesis 4

	CGPA-group	N	Mean	Std. Deviation	Std. Error Mean
School District B Scores	2.5 and above	37	222.4054	16.32050	2.68307
	below 2.5	13	196.9231	10.27506	2.84979

Figure 34 indicates there is a noticeable difference in means and range of EOC scores based on CGPA.

Figure 34 – Box Plot, School District B, PPT, No IEP, Relationship of CGPA and EOC Scale Scores



Null Hypothesis 5f: For non-IEP students taking the United States Government EOC examination by computer-based mode of delivery, non-IEP students with a cumulative GPA of 2.5 or higher will not score higher than non-IEP students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 44, the difference between the two CGPA groups is observable. The researcher rejected the null hypothesis and data supports the alternative [$t(46) = 6.407$, $p < 0.0005$] indicating there is a contribution to a relationship between students without IEPs, CGPA, and EOC scores when taking a CBT type of EOC exam. The EOC scale scores for students without IEPs at school A, who had CGPAs of 2.5 and above, are significantly greater than the EOC scale scores of students without IEPs with CGPAs below 2.5.

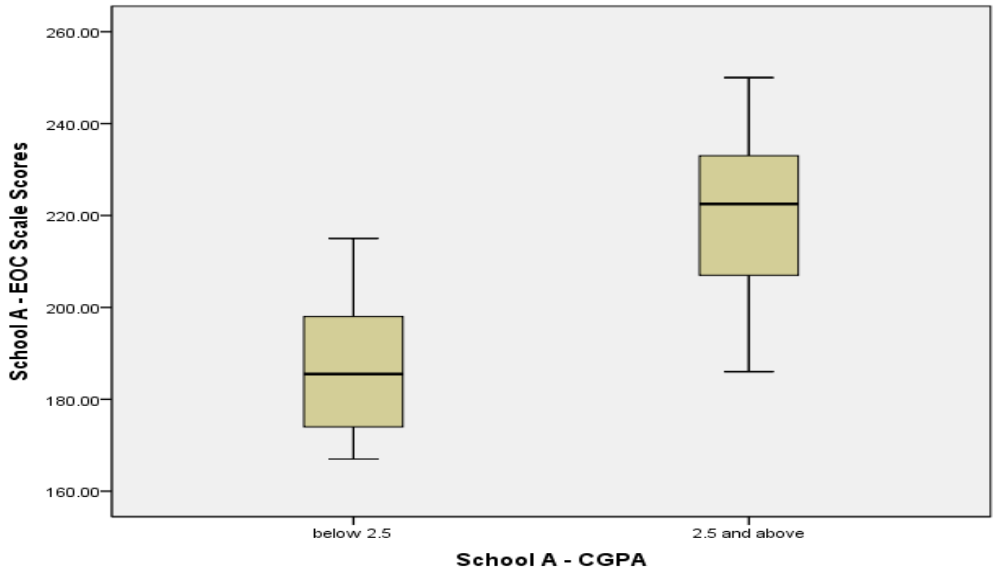
Table 44

Statistics for School District A, CBT, IEP, Null Hypothesis 5

	GPA-group	N	Mean	Std. Deviation	Std. Error Mean
EOC	2.5 and above	32	220.3125	18.45559	3.26252
	below 2.5	16	186.2500	14.85261	3.71315

According to Figure 35, the difference in means supports the alternative hypothesis that there is a relationship for students without IEPs, between students' CGPA and EOC scores, when taking a CBT form of the EOC exam.

Figure 35 – Box Plot, School District A, CBT, No IEP, Relationship of CGPA and EOC Scale Scores



Statistical Analysis, Ethnicity, Black

Null Hypothesis 1g: For Black students taking the United States Government EOC examination by paper-pencil mode of delivery, compared to Black students taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam.

According to the descriptive statistics in Table 45, the EOC scale scores for black students at schools A and B show there is a small observable difference between the EOC scale scores for black students at schools A and B. Application of the *t*-test for difference in means [t (38) = 1.251, p = 0.219] allowed the researcher to not reject the null hypothesis that there were no differences in student test scores based on mode of test administration.

Table 45

Descriptive Statistics Comparing Black Students, Schools A and B

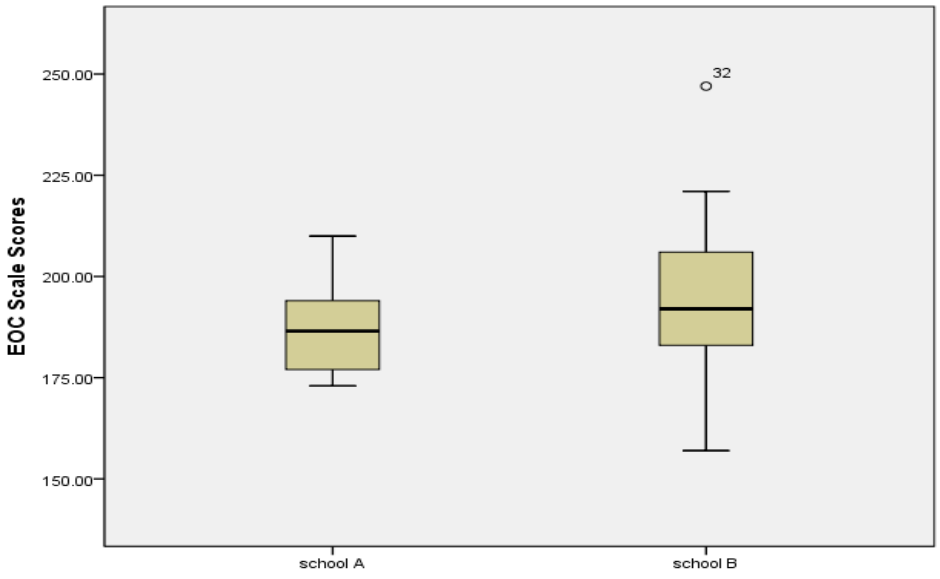
		N	Mean	Std. Deviation	Std. Error Mean
EOC	school A	20	187.9000	11.77821	2.63369
	school B	20	194.7500	21.47183	4.80125

Note: School A used computer-based and School B used paper-pencil-based testing.

The box plot in Figure 36 demonstrates there is very little difference in means and ranges of EOC Scale Scores for black students at Schools A and B

Figure 36 – Box Plot, Comparison Black EOC Scale Scores by School District

School District A = CBT School District B = PPT



Null Hypothesis 2g: For Black students taking the United States Government

EOC examination by paper-pencil mode of delivery, there is no relationship between with a Black student cumulative GPA and United States Government EOC exam score.

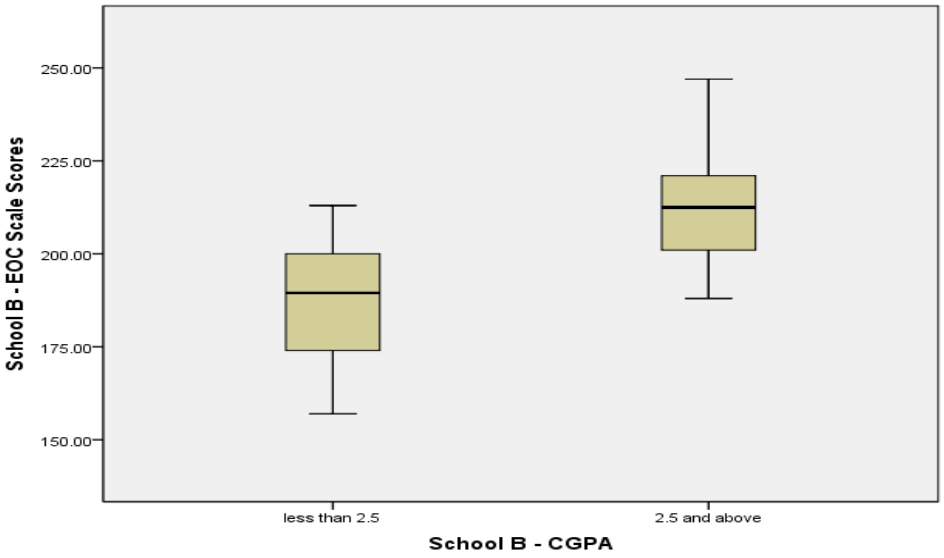
According to Table 46, the Pearson Product Moment Correlation Coefficient demonstrates a moderately strong relationship that is statistically significant [$R(18) = 0.579$, $p = 0.008$]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between black students' CGPA and EOC scores when taking the PPT form of the U.S. Government EOC.

Table 46

Pearson Product Moment Correlation Coefficient

		EOC scores-School B	CGPA-School B
EOC scores-School B	Pearson Correlation	1	.579**
	Sig. (2-tailed)		.008
	N	20	20
CGPA-School B	Pearson Correlation	.579**	1
	Sig. (2-tailed)	.008	
	N	*20*	*20*

Figure 37 Box Plot, School District B, PPT, Black, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 3g: For Black students taking the United States Government EOC examination by computer-based mode of delivery, there is no relationship between a Black student cumulative GPA and United States Government EOC exam score.

According to Table 47 the Pearson Product Moment Correlation Coefficient demonstrates a strong correlation that is statistically significant [$R(18) = 0..664$, $p = 0.001$,]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between black students' CGPA and EOC scores when taking a CBT type of EOC exam.

Table 47

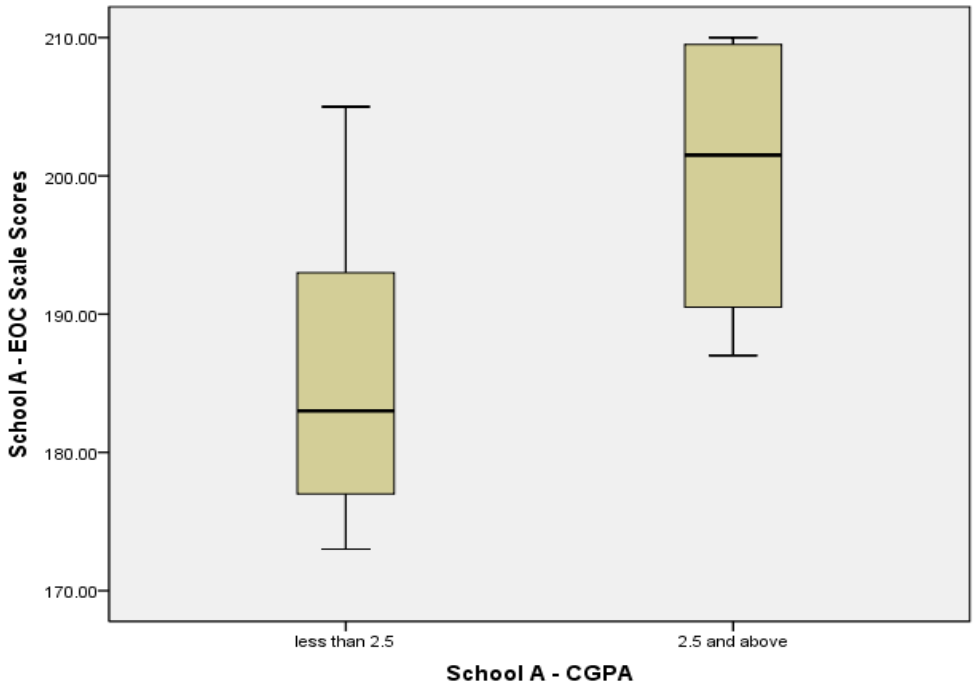
Pearson Product Moment Correlation Coefficient

		EOC	CGPA
EOC	Pearson Correlation	1	.664**
	Sig. (2-tailed)		.001
	N	20	20
CGPA	Pearson Correlation	.664**	1
	Sig. (2-tailed)	.001	
	N	20	20

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 38 demonstrates a significant difference in means and ranges of scores between below 2.5, and 2.5 and above CGPAs.

Figure 38 – Box Plot, School District A, CBT, Black, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 4g: For Black students taking the United States Government EOC examination by paper-pencil mode of delivery, Black students with a cumulative GPA of 2.5 or higher will not score higher than Black students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 48, there is an observable difference between the means for the two samples. A Chi-Square test for Homogeneity applied to data from the PPT administration of the EOC resulted in [X (20) = 0 and X-critical = 3.845] which allowed the researcher to not reject the null hypothesis. A Confidence Interval test for difference in means resulted in the range [-163.6, 217.5]. Since the test value of 0 is in the range, the researcher did not reject the null hypothesis. Data from this study indicates that there is no significant difference between Black students with CGPAs of 2.5 or above and those with a CGPA of less than 2.5 on PPT scores on Government EOC exams.

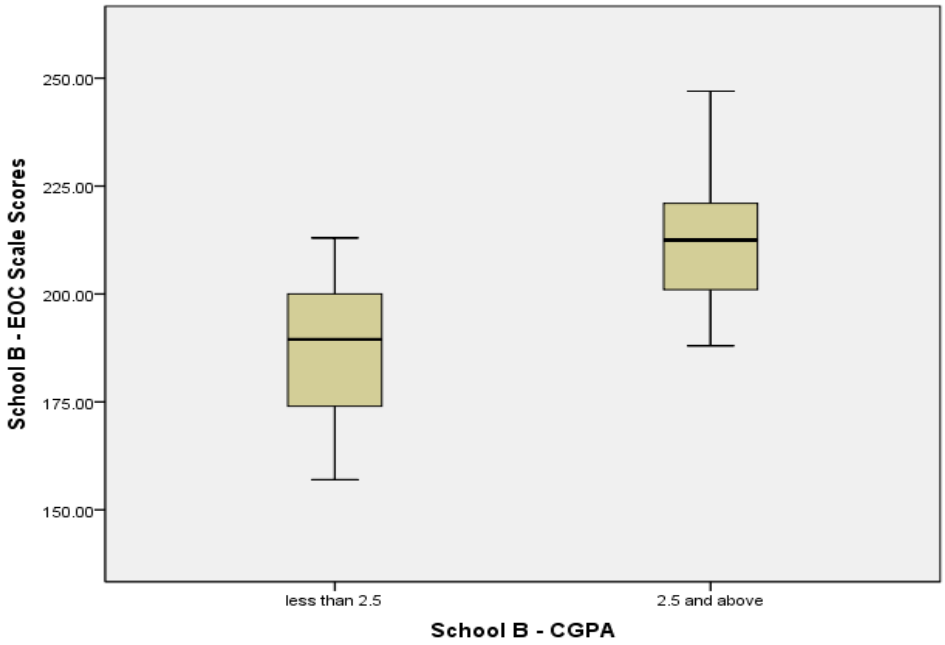
Table 48

Statistics for School District B, PPT, Black, Null Hypothesis 4

	CGPA-group	N	Mean	Std. Deviation	Std. Error Mean
School District B Scores	2.5 and above	6	213.6667	20.49065	8.36527
	less than 2.5	14	186.6429	16.62532	4.44330

Figure 39 indicates there is an observable difference in means and ranges of EOC scores based on CGPA.

Figure 39 – Box Plot, School District B, PPT, Black, Relationship of CGPA and EOC Scale Scores



Null Hypothesis 5g: For Black students taking the United States Government EOC examination by computer-based mode of delivery, Black students with a cumulative GPA of 2.5 or higher will not score higher than Black students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to Table 49, there is an observable difference between the means for the two samples. A Chi-Square test for Homogeneity applied to data from the CBT administration of the EOC resulted in [X (20) = 0 and X-critical = 3.845] which allowed

the researcher to not reject the null hypothesis. A Confidence Interval test for difference in means resulted in the range [-203.0, 233.3]. Since the test value of 0 is in the range, the researcher did not reject the null hypothesis. Data from this study indicates that there is no significant difference between Black students with CGPAs of 2.5 or above and those with a CGPA of less than 2.5 on CBT scores on Government EOC exams.

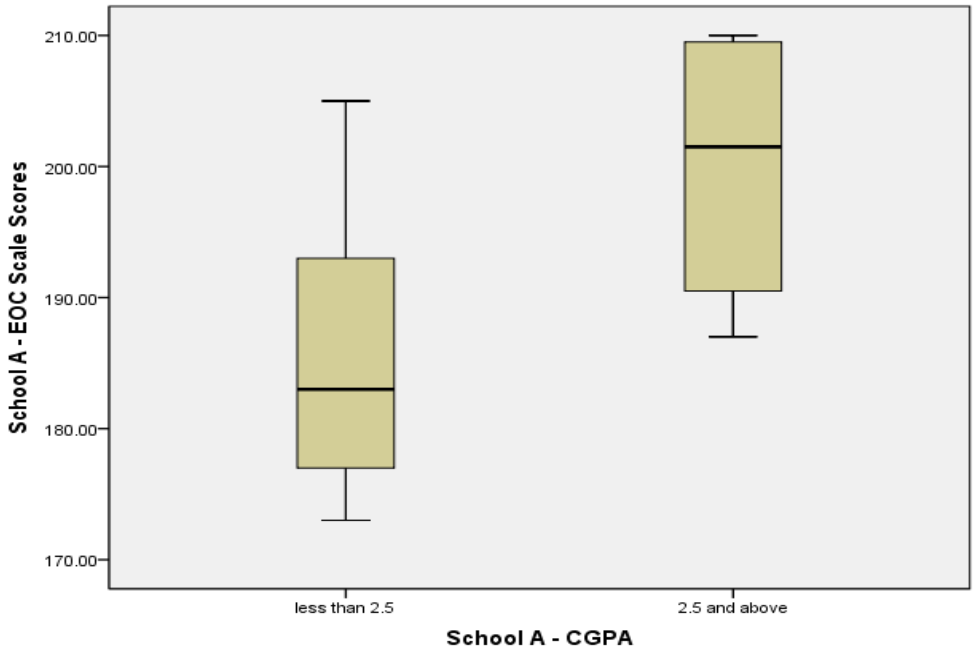
Table 49

Statistics for School District A, CBT, Black, Null Hypothesis 5

	GPA-group	N	Mean	Std. Deviation	Std. Error Mean
EOC	2.5 and above	4	200.0000	11.34313	5.67157
	less than 2.5	16	184.8750	10.05899	2.51475

According to Figure 40, the difference in means supports the alternative hypothesis that there is a relationship for Black students, between students' CGPA and EOC scores, when taking a CBT form of the EOC exam.

Figure 40 – Box Plot Showing Relationship of CGPA and EOC Scale Scores For Black students in School District A, When Taking CBT



Statistical Analysis, Ethnicity, Asian

Null Hypothesis 1h: For Asian students taking the United States Government EOC examination by paper-pencil mode of delivery, compared to Asian students taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam.

According to the descriptive statistics in Table 50, the EOC scale scores for Asian students at school B are greater than the EOC scale scores of Asian students at school A. The difference is significant, which was supported by the *t*-test [t (30.96) = 2.386, p = 0.023]: therefore the researcher rejected the null hypothesis and did not reject the alternative hypothesis that there are differences in student test scores based on mode of test administration.

Table 50

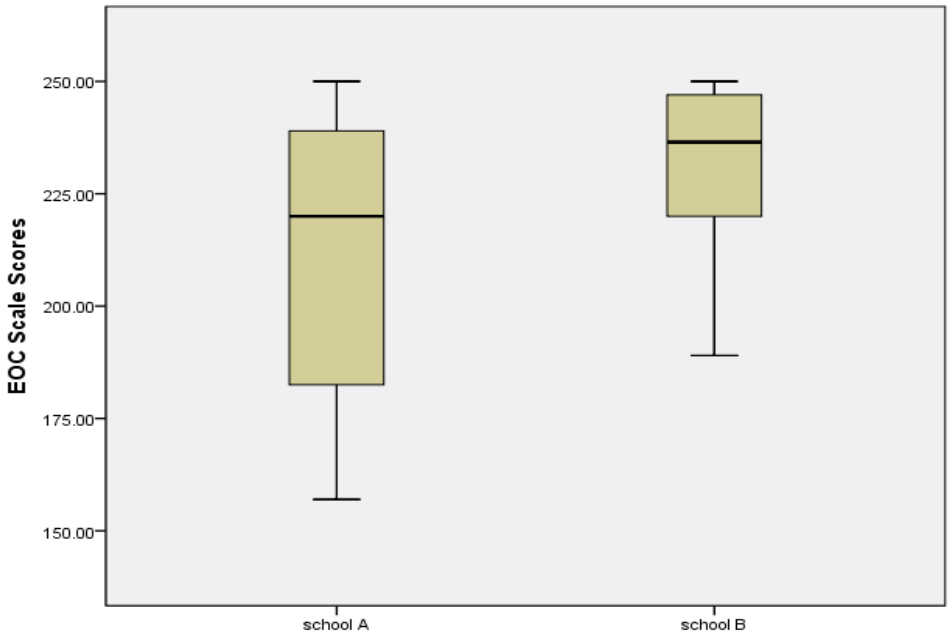
Descriptive Statistics Comparing Asian Students, Schools A and B

	N	Mean	Std. Deviation	Std. Error Mean
EOC school A	20	212.1500	30.60319	6.84308
EOC school B	20	231.1500	18.21588	4.07320

Note: School A used computer-based and School B used paper-pencil-based testing.

The box plot in Figure 41 demonstrates there is a difference in means and ranges of EOC Scale Scores for Asian students at Schools A and B.

Figure 41 – Box Plot, Comparison Asian EOC Scale Scores by School District
School District A = CBT School District B = PPT



Null Hypothesis 2h: For Asian students taking the United States Government EOC examination by paper-pencil mode of delivery, there is no relationship between an Asian student cumulative GPA and United States Government EOC exam score.

According to Table 51, the Pearson Product Moment Correlation Coefficient demonstrates a strong relationship that is statistically significant [$R(18) = 0.635$, $p = 0.003$]: therefore the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between Asian students' CGPA and EOC scores when taking the PPT form of the U.S. Government EOC.

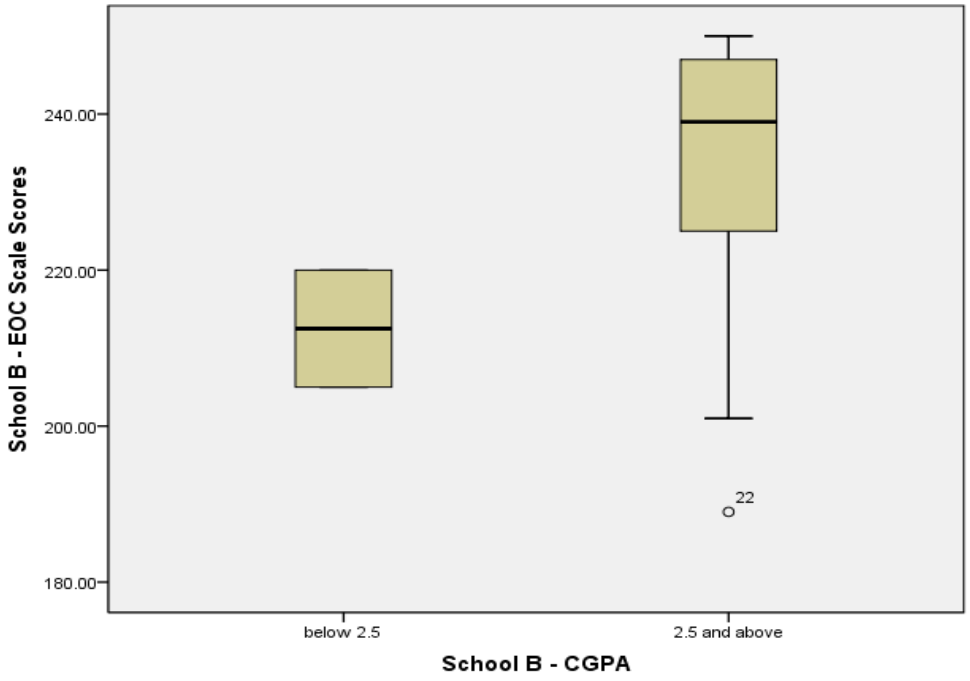
Table 51

Pearson Product Moment Correlation Coefficient

		EOC scores-School B	CGPA-School B
EOC scores-School B	Pearson Correlation	1	.635**
	Sig. (2-tailed)		.003
	N	20	20
CGPA-School B	Pearson Correlation	.635**	1
	Sig. (2-tailed)	.003	
	N	20	20

**Correlation significant at the 0.01 level (2-tailed).

Figure 42 Box Plot, School District B, PPT, Black, Correlation between CGPA And EOC Scale Scores.



Null Hypothesis 3h: For Asian students taking the United States Government EOC examination by computer-based mode of delivery, there is no relationship between with an Asian student cumulative GPA and United States Government EOC exam score.

According to Table 52, the Pearson Product Moment Correlation Coefficient demonstrates a strong correlation that is statistically significant [$R(18) = 0.846$, $p < 0.0005$]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between Asian students' CGPA and EOC scores when taking a CBT.

Table 52

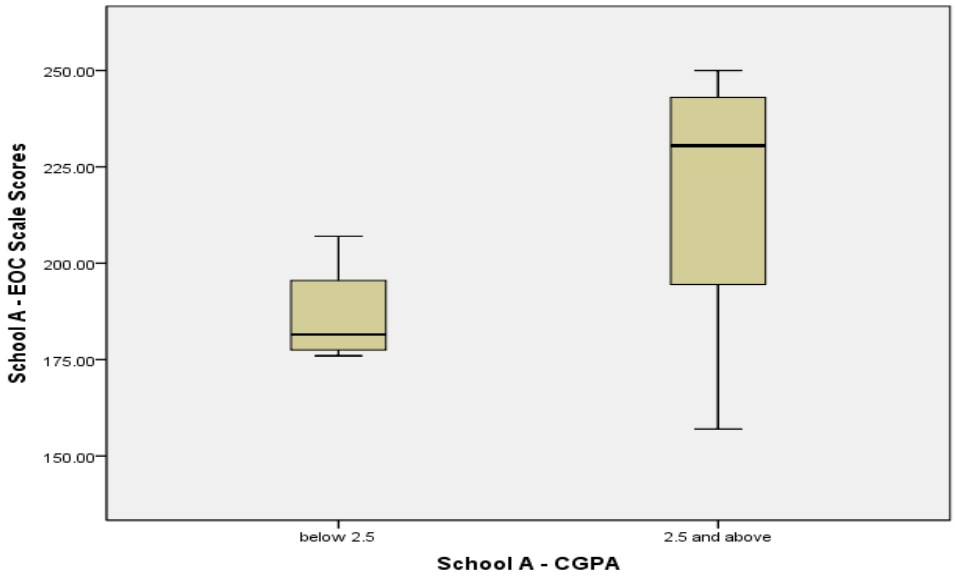
Pearson Product Moment Correlation Coefficient

		EOC	CGPA
EOC	Pearson Correlation	1	.846**
	Sig. (2-tailed)		.000
	N	20	20
CGPA	Pearson Correlation	.846**	1
	Sig. (2-tailed)	.000	
	N	20	20

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 43 demonstrates a significant difference in means and ranges of scores between below 2.5, and 2.5 and above CGPAs.

Figure 43 – Box Plot, School District A, CBT, Black, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 4h: For Asian students taking the United States Government EOC examination by paper-pencil mode of delivery, Asian students with a cumulative GPA of 2.5 or higher will not score higher than Asian students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to the descriptive statistics in Table 53, there is an observable difference between the EOC scale scores for Asian students at school B. Application of a *t*-test for difference in means [$t(18) = 1.586$, $p = 0.130$] allowed the researcher to not reject the null hypothesis. These results were verified with a Chi Square Test for Homogeneity [X=0; X-critical = 3.845] and Confidence Interval Tests for Difference in Means [-1304, 1406].

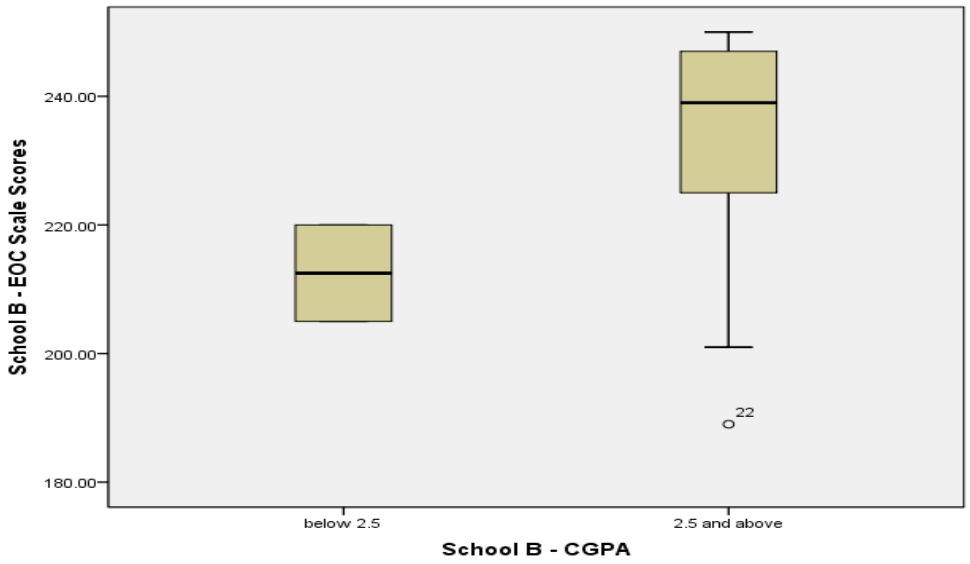
Table 53

Statistics for School District B, PPT, Asian, Null Hypothesis 4

	CGPA-group	N	Mean	Std. Deviation	Std. Error Mean
School District B Scores	2.5 and above	18	233.2222	17.85417	4.20827
	below 2.5	2	212.5000	10.60660	7.50000

Figure 44 indicates there is an observable difference in means and ranges of EOC scores based on CGPA.

Figure 44 – Box Plot, School District B, PPT, Asian, Relationship of CGPA and EOC Scale Scores



Null Hypothesis 5h: For Asian students taking the United States Government EOC examination by computer-based mode of delivery, Asian students with a cumulative GPA of 2.5 or higher will not score higher than Asian students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to the descriptive statistics in Table 54, there is an observable difference between the EOC scale scores for Asian students at school A. Results of a *t*-test for difference in means [$t(18) = 2.020$, $p = 0.058$] allowed the researcher to not reject the null hypothesis. These results were verified with a Chi Square Test for Homogeneity [X=0; X-critical = 3.845] and Confidence Interval Tests for Difference in Means [-703.6, 893.5].

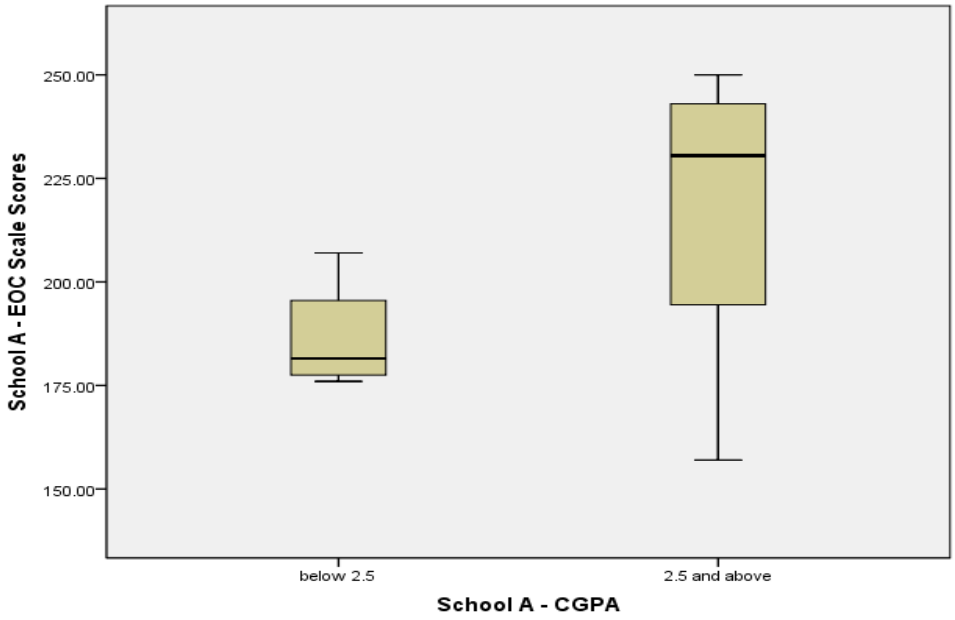
Table 54

Statistics for School District A, CBT, Asian, Null Hypothesis 5

CGPA-group	N	Mean	Std. Deviation	Std. Error Mean
EOC 2.5 and above	16	218.5625	30.45427	7.61357
EOC below 2.5	4	186.5000	14.05940	7.02970

According to Figure 45, the observed difference in means supports the null hypothesis.

Figure 45 – Box Plot Showing Relationship of CGPA and EOC Scale Scores For Asian Students in School District A, When Taking CBT



Statistical Analysis, Ethnicity, White

Null Hypothesis 1i: For White students taking the United States Government EOC examination by paper-pencil mode of delivery, compared to White students taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam.

According to the descriptive statistics in Table 55, there is a small observable difference between EOC scale scores for White students at schools A and B. Application of a *t*-test for difference in means [t (38) = 1.691, p = 0.099] allowed the researcher to reject the null hypothesis.

Table 55

Descriptive Statistics Comparing White Students, Schools A and B

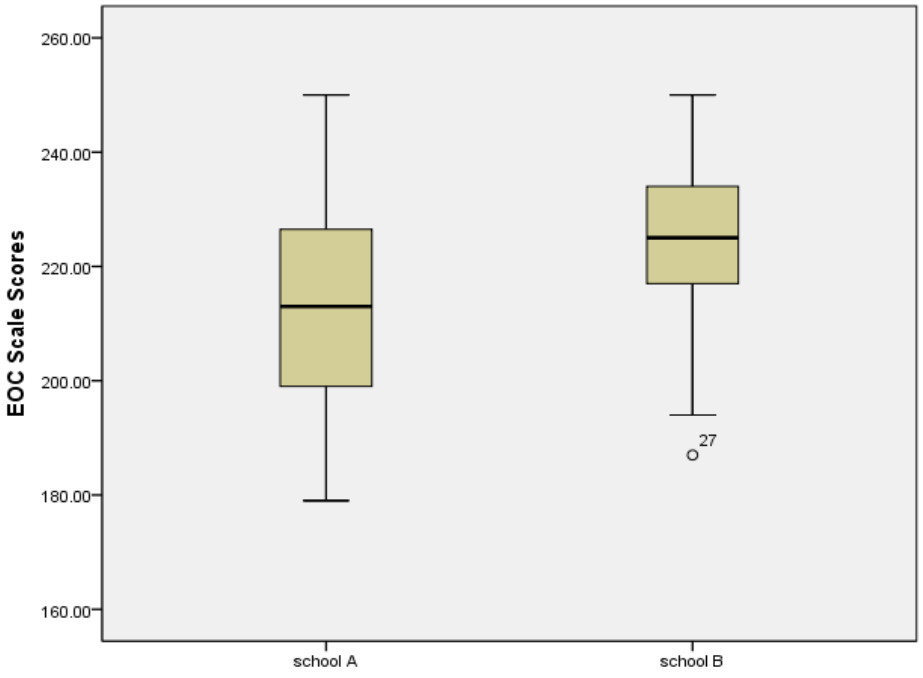
	N	Mean	Std. Deviation	Std. Error Mean
EOC school A	20	214.4500	20.10623	4.49589
EOC school B	20	224.2000	16.15256	3.61182

Note: School A used computer-based and School B used paper-pencil-based testing.

The box plot in Figure 46 demonstrates there is a difference in means and ranges of EOC Scale Scores for White students at Schools A and B.

Figure 46 – Box Plot, Comparison White EOC Scale Scores by School District

School District A = CBT School District B = PPT



Null Hypothesis 2i: For White students taking the United States Government EOC examination by paper-pencil mode of delivery, there is no relationship between a White student cumulative GPA and United States Government EOC exam score.

According to Table 56, the Pearson Product Moment Correlation Coefficient demonstrates there is a weak relationship that is not significant [$R(18) = 0.395$, $p =$

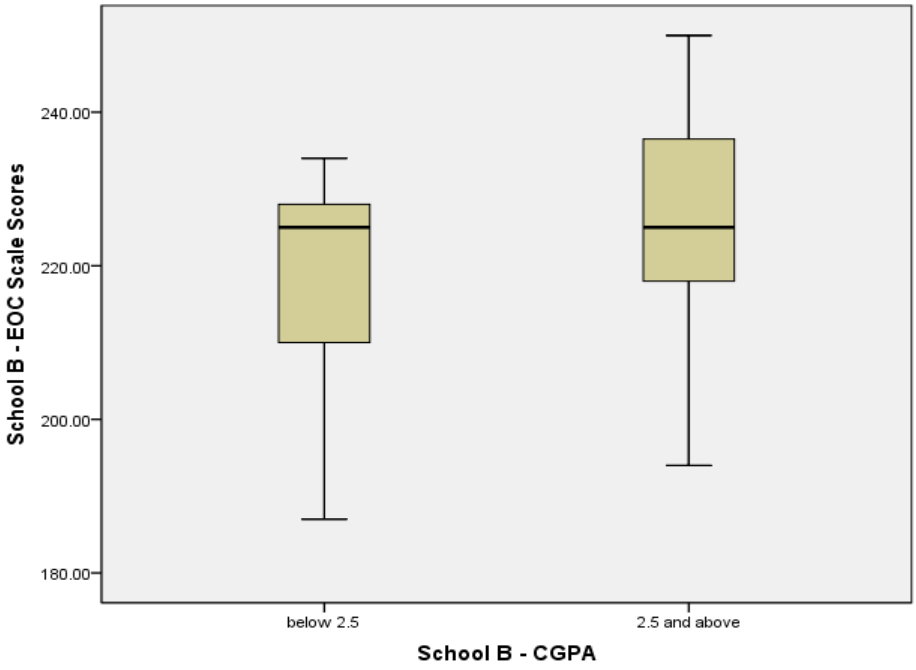
0.085]: therefore, the researcher did not reject the null hypothesis.

Table 56

Pearson Product Moment Correlation Coefficient

		EOC scores-School B	CGPA-School B
EOC scores-School B	Pearson Correlation	1	.395
	Sig. (2-tailed)		.085
	N	20	20
CGPA-School B	Pearson Correlation	.395	1
	Sig. (2-tailed)	.085	
	N	20	20

Figure 47 Box Plot, School District B, PPT, White, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 3i: For White students taking the United States Government EOC examination by computer-based mode of delivery, there is no relationship between with a White student cumulative GPA and United States Government EOC exam score.

According to Table 57, the Pearson Product Moment Correlation Coefficient demonstrates a strong relationship that is statistically significant [$R(18) = 0.660$, $p =$

0.002]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between White students' CGPA and EOC scores when taking a CBT type EOC exam.

Table 57

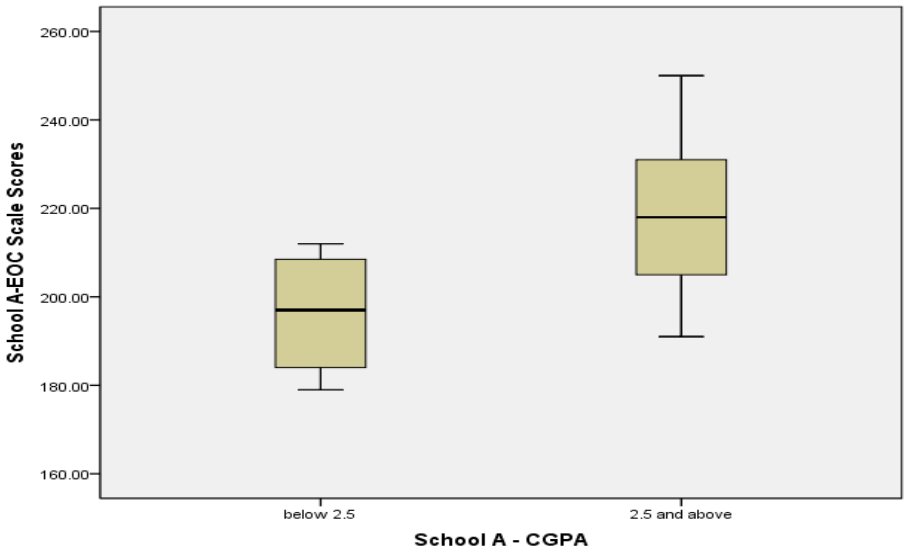
Pearson Product Moment Correlation Coefficient

		EOC	CGPA
EOC	Pearson Correlation	1	.660**
	Sig. (2-tailed)		.002
	N	20	20
CGPA	Pearson Correlation	.660**	1
	Sig. (2-tailed)	.002	
	N	20	20

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 48 demonstrates a significant difference in means and ranges of scores between below 2.5, and 2.5 and above CGPAs.

Figure 48 – Box Plot, School District A, CBT, White, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 4i: For White students taking the United States Government EOC examination by paper-pencil mode of delivery, White students with a cumulative GPA of

2.5 or higher will not score higher than White students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to the descriptive statistics in Table 58, there is a small observable difference between the EOC scale scores for White students at school B. Application of a *t*-test for difference in means [$t(18) = 1.196$, $p = 0.247$] allowed the researcher to not reject the null hypothesis. These results were verified with a Chi Square Test for Homogeneity [X=0; X-critical = 3.845] and Confidence Interval Tests for Difference in Means [-457.7, 477.2].

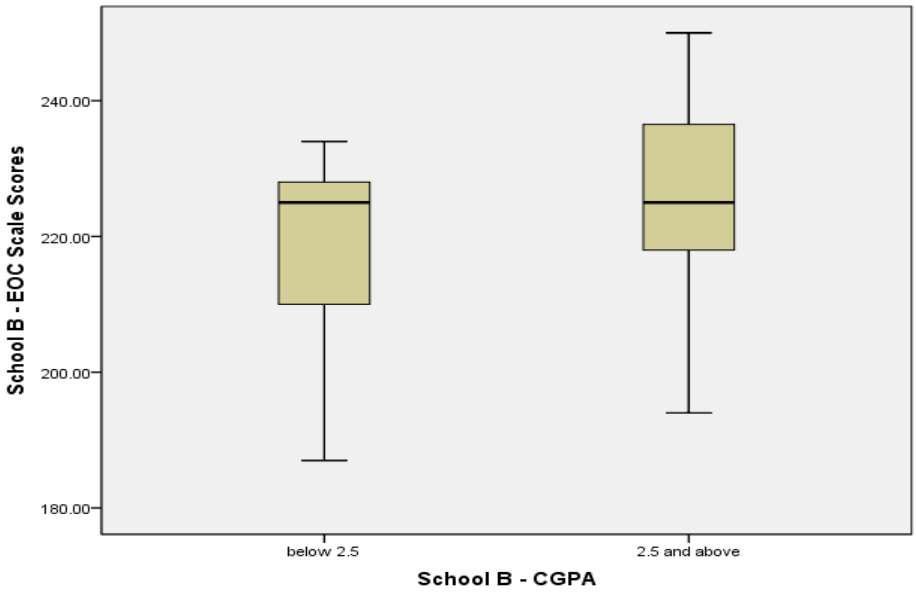
Table 58

Statistics for School District B, PPT, White, Null Hypothesis 4

	CGPA-group	N	Mean	Std. Deviation	Std. Error Mean
School District B Scores	2.5 and above	15	226.6667	15.04596	3.88485
	below 2.5	5	216.8000	18.86001	8.43445

Figure 49 indicates there is no observable difference in means based on CGPA.

Figure 49 – Box Plot, School District B, PPT, White, Relationship of CGPA and EOC Scale Scores



Null Hypothesis 5i: For White students taking the United States Government EOC examination by computer-based mode of delivery, White students with a cumulative GPA of 2.5 or higher will not score higher than White students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to the descriptive statistics in Table 59, the EOC scale scores of White students at school A who had CGPAs of 2.5 and above are greater than the EOC scale scores of students who had CGPAs below 2.5. The *t*-test for difference in means results [$t(18) = 2.224$, $p = 0.039$] allowed the researcher to not reject the null hypothesis. These results were verified with a Chi Square Test for Homogeneity [X=0; X-critical = 3.845] and Confidence Interval Tests for Difference in Means [-428.7, 474.5].

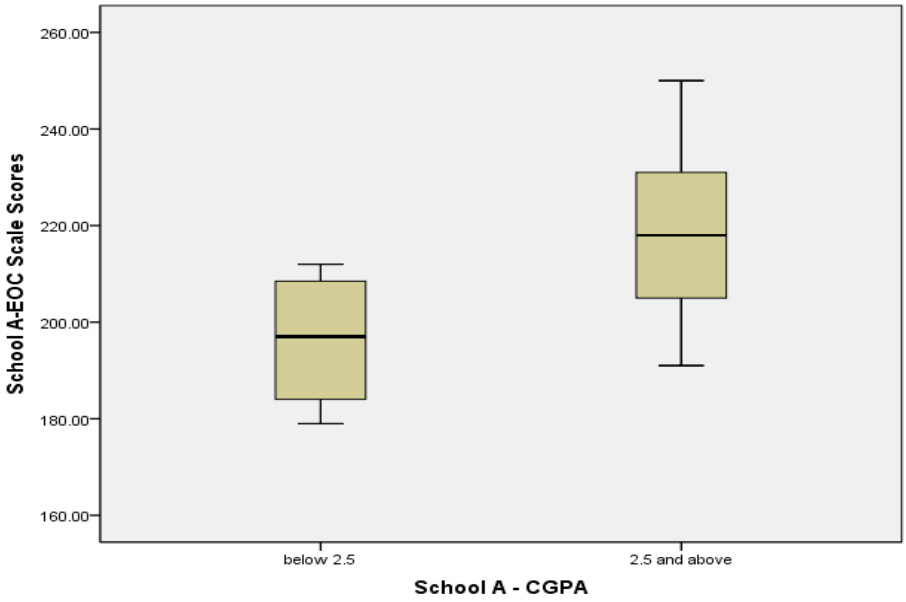
Table 59

Statistics for School District A, CBT, White, Null Hypothesis 5

	GPA-group	N	Mean	Std. Deviation	Std. Error Mean
EOC	2.5 and above	16	219.0000	18.88562	4.72141
	below 2.5	4	196.2500	14.99722	7.49861

According to Figure 50, the observed difference in means and ranges supports the null hypothesis.

Figure 50 – Box Plot, School District A, CBT, White, Relationship of CGPA and EOC Scale Scores



Statistical Analysis, Ethnicity, Hispanic

Null Hypothesis 1j: For Hispanic students taking the United States Government EOC examination by paper-pencil mode of delivery, compared to Hispanic students taking the examination in computer-based mode of delivery, there is no difference in average students' scores on the United States Government EOC exam.

According to the descriptive statistics in Table 60, the EOC scale scores of Hispanic students at school B are greater than the EOC scale scores of Hispanic students at school A. The *t*-test for difference in means [t (38) = 2.346, p = 0.024] allowed the researcher to reject the null hypothesis. .

Table 60

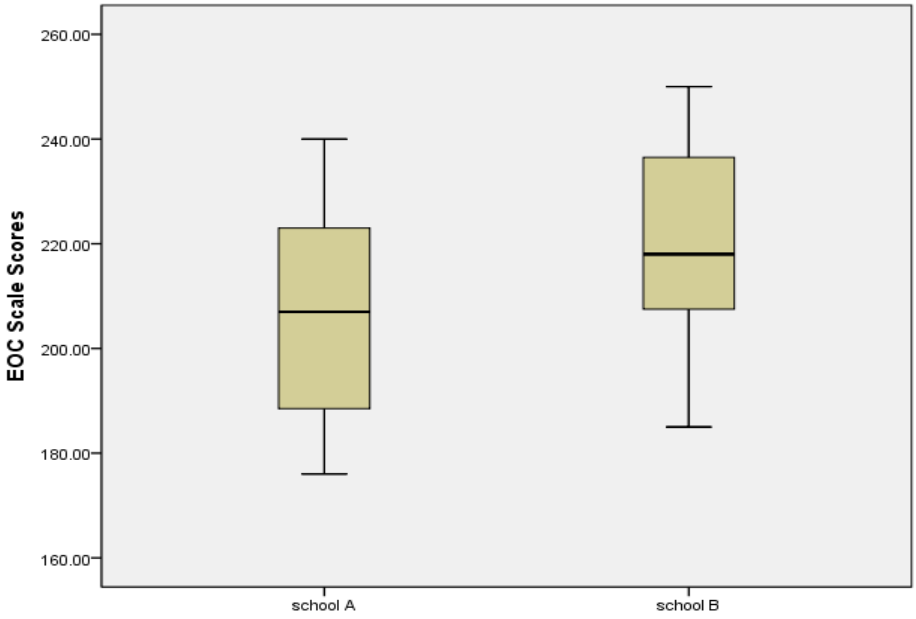
Descriptive Statistics Comparing Hispanic Students, Schools A and B

		N	Mean	Std. Deviation	Std. Error Mean
EOC	school A	20	206.5000	19.71775	4.40902
	school B	20	220.6500	18.41131	4.11689

Note: School A used computer-based and School B used paper-pencil-based testing.

The box plot in Figure 51 demonstrates there is a difference in means and ranges of EOC Scale Scores for Hispanic students at Schools A and B.

Figure 51 – Box Plot, Comparison Hispanic EOC Scale Scores by School District
School District A = CBT School District B = PPT



Null Hypothesis 2j: For Hispanic students taking the United States Government EOC examination by paper-pencil mode of delivery, there is no relationship between with a Hispanic student cumulative GPA and United States Government EOC exam score.

According to Table 61, the Pearson Product Moment Correlation Coefficient demonstrates a weak, significant, relationship [$R(18) = 0.488$, $p = 0.029$]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a weak positive relationship between female students CGPA and EOC scores for Hispanic students taking the PPT form of the U.S. Government EOC.

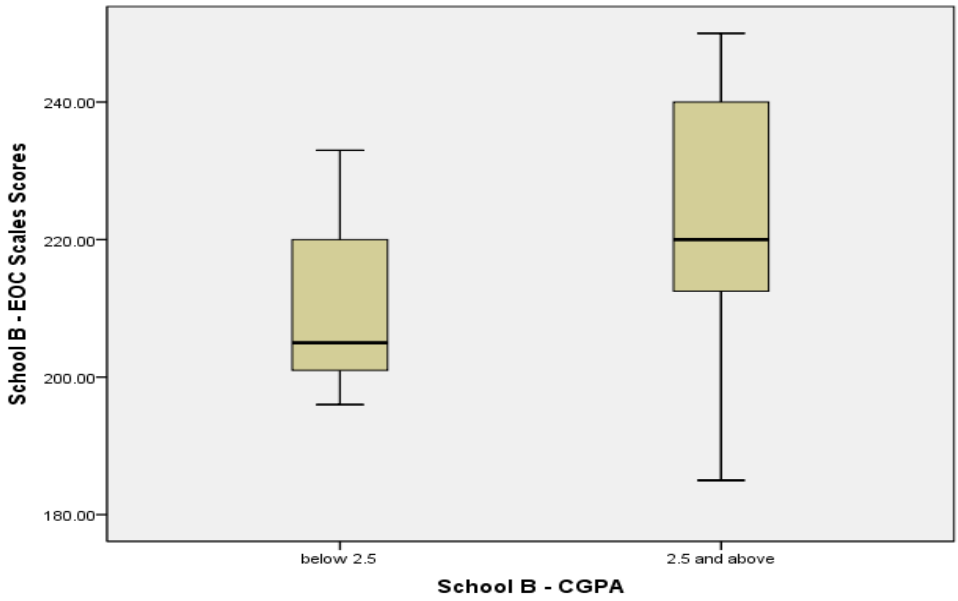
Table 61

Pearson Product Moment Correlation Coefficient

		EOC scores-School B	CGPA-School B
EOC scores-School B	Pearson Correlation	1	.488*
	Sig. (2-tailed)		.029
	N	20	20
CGPA-School B	Pearson Correlation	.488*	1
	Sig. (2-tailed)	.029	
	N	20	20

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 52 Box Plot, School District B, PPT, Hispanic, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 3j: For Hispanic students taking the United States Government EOC examination by computer-based mode of delivery, there is no relationship between with a Hispanic student cumulative GPA and United States Government EOC exam score.

According to Table 62, the Pearson Product Moment Correlation Coefficient demonstrates a strong relationship that is statistically significant [$R(16) = 0.625$, $p =$

0.006]: therefore, the researcher rejected the null hypothesis and data supports the alternative, indicating there is a positive relationship between Hispanic students CGPA and EOC scores when taking a CBT type of EOC exam.

Table 62

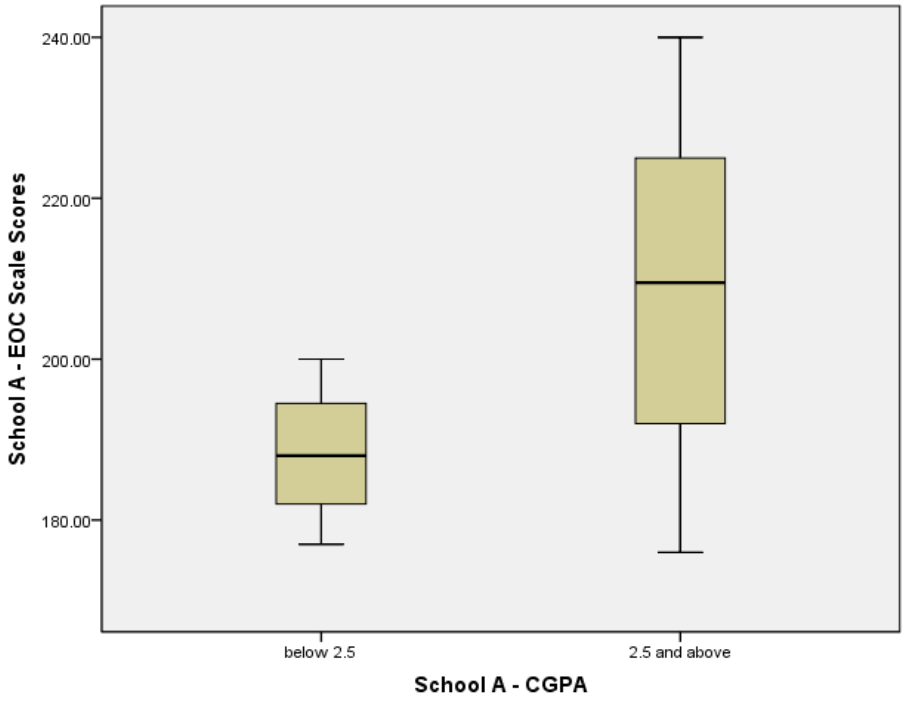
Pearson Product Moment Correlation Coefficient

		EOC	CGPA
EOC	Pearson Correlation	1	.625**
	Sig. (2-tailed)		.006
	N	20	18
CGPA	Pearson Correlation	.625**	1
	Sig. (2-tailed)	.006	
	N	18	18

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 53 demonstrates a significant difference in means and the ranges of scores between below 2.5, and 2.5 and above CGPAs.

Figure 53 – Box Plot, School District A, CBT, Hispanic, Correlation between CGPA and EOC Scale Scores



Null Hypothesis 4j: For Hispanic students taking the United States Government EOC examination by paper-pencil mode of delivery, Hispanic students with a cumulative GPA of 2.5 or higher will not score higher than Hispanic students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to the descriptive statistics in Table 63, there is a small observable difference between the EOC scale scores for Hispanic students at school B. Application of a *t*-test for difference in means [$t(18) = 1.386$, $p = 0.183$] allowed the researcher to not reject the null hypothesis. These results were verified with a Chi Square Test for Homogeneity [X=0; X-critical = 3.845] and Confidence Interval Tests for Difference in Means [-366.4, 391.9].

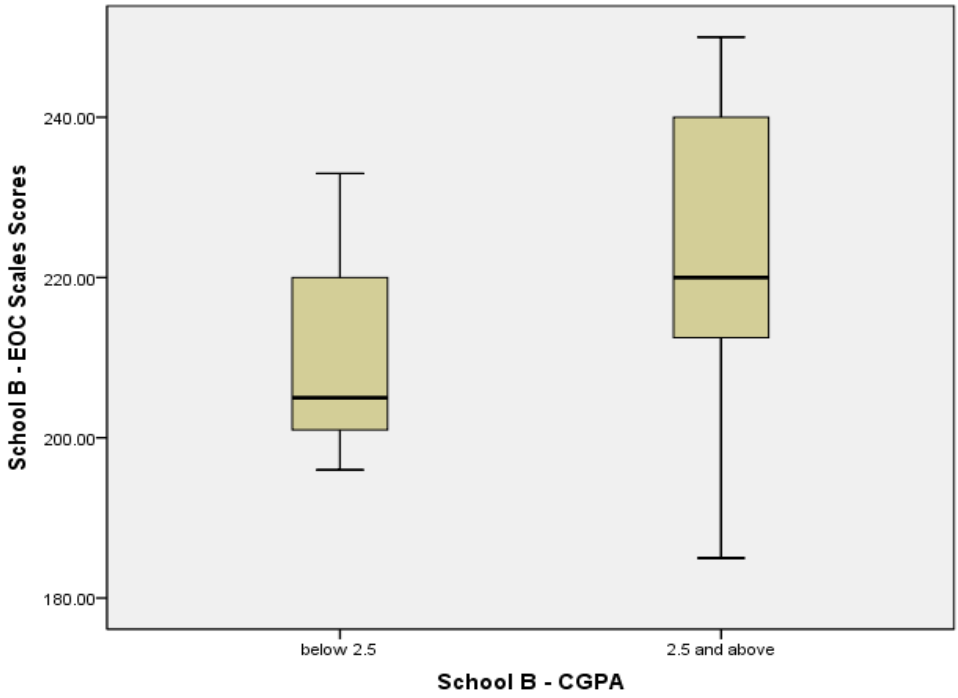
Table 63

Statistics for School District B, PPT, Hispanic, Null Hypothesis 4

	CGPA-group	N	Mean	Std. Deviation	Std. Error Mean
School District B Scores	2.5 and above	15	223.8667	18.69632	4.82737
	below 2.5	5	211.0000	15.21512	6.80441

Figure 54 demonstrates a difference in means and ranges of scores based on CGPA.

Figure 54 – Box Plot, School District B, PPT, Hispanic, Relationship of CGPA and EOC Scale Scores



Null Hypothesis 5j: For Hispanic students taking the United States Government EOC examination by computer-based mode of delivery, Hispanic students with a cumulative GPA of 2.5 or higher will not score higher than Hispanic students with a cumulative GPA of less than 2.5, on the United States Government EOC exam.

According to the descriptive statistics in Table 64, there is an observable difference between the EOC scale scores of Hispanic students at school A based on CGPA. An application of the *t*-test for difference in means [t (16) = 2.063, p = 0.056] allowed the researcher to not reject the null hypothesis. These results were verified with a Chi Square Test for Homogeneity [X=0; X-critical = 3.845] and Confidence Interval Tests for Difference in Means [-303.3, 343.1].

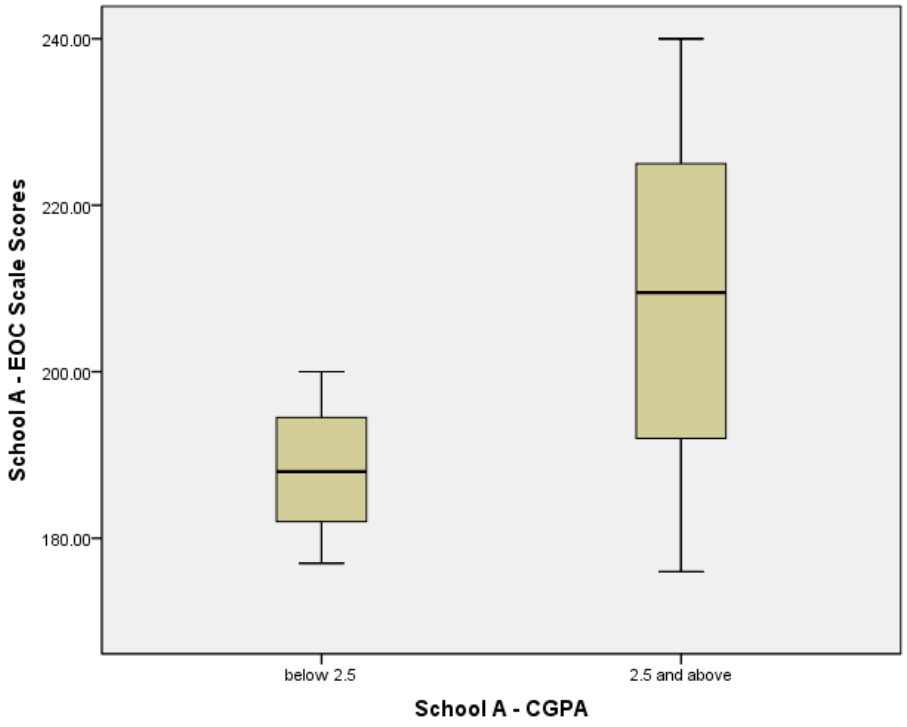
Table 64

Statistics for School District A, CBT, White, Null Hypothesis 5

	CGPA-group	N	Mean	Std. Deviation	Std. Error Mean
Score	2.5 and above	14	208.4286	18.59989	4.97103
	below 2.5	4	188.2500	9.42956	4.71478

According to Figure 55, the observed difference in means and ranges supports the null hypothesis.

Figure 55 – Box Plot, School District A, CBT, Hispanic, Relationship of CGPA and EOC Scale Scores



Summary

Table 65

Overall Findings

Null Hypothesis	Overall Group	Gender		FRL		IEP		Ethnicity			
		F	M	Yes	No	Yes	No	B	A	W	H
1	NR	R	NR	NR	NR	NR	NR	NR	R	R	R
2	R	R	R	R	R	R	R	R	R	NR	R
3	R	R	R	R	R	R	R	R	R	R	R
4	NR	NR	R	NR	NR	R	NR	NR	NR	NR	NR
5	NR	NR	R	NR	R	R	NR	NR	NR	NR	NR

Note: R = reject null hypothesis; NR = does not reject null hypothesis

Table 66

Pearson Product Moment Correlation Coefficients for Hypotheses 2 & 3

NCLB Subgroups	Hypothesis 2 PPT/GPA	Hypothesis 3 CBT/GPA
Overall	.719	.752
Female	.682	.568
Male	.615	.644
Yes FRL	.336	.512
No FRL	.552	.540
Yes IEP	.553	.461
No IEP	.664	.747
Black	.579	.664
Asian	.635	.846
White	.395	.660
Hispanic	.488	.625

In conclusion, females, Asians, Whites, and Hispanics demonstrated a difference in average students' scores between CBT and PPT. However, only Whites taking PPT had no relationship between student CGPA and EOC scores. In general both PPT and CBT

students' scores were higher on the EOC when they have a higher CGPA, as expected. Both overall and all subgroups demonstrated there was a relationship between student CGPA and EOC scores with CBT. Overall and for all subgroups except males and IEP students taking the EOC by PPT, demonstrated that students with a cumulative GPA of 2.5 or higher will not score higher than students with a cumulative GPA of less than 2.5, on the United States Government EOC exam. Overall and for all subgroups except males, non-FRL, and those with IEPs taking the EOC by CBT demonstrated students with a cumulative GPA below 2.5 will not score higher than students with a cumulative GPA of 2.5 or above, on the United States Government EOC exam.

Looking at the overall data, Null Hypothesis 1 was not rejected demonstrating there is no difference in average students' scores on the United States Government EOC, whether the student took the test PPT or CBT. Null Hypotheses 2 and 3 were rejected, indicating that for both PPT and CBT there is a relationship between student cumulative GPA and United States Government EOC exam scores. The data that was presented in this chapter, particularly Null Hypotheses 4 and 5 demonstrates there is little relationship between CGPA and EOC scores depending upon the mode of test administration, except as mentioned above in a few subgroups. Further discussion regarding these findings is found in Chapter 5.

Chapter Five: Implications and Recommendations

Overview

Technology is changing the slant of education throughout the world. The current generation moving through schools is tech savvy. These students grew up in a world saturated with electronic equipment such as video games, cell phones, laptops, iPods, Kindles, mp3 players, and many other devices that make them part of the technophile generation. The sources they have had access to, which have helped socialize them since birth, include such elements as Facebook, Myspace, Twitter, email, and assorted internet sites. From the home environment, to the work place, to the education setting students are utilizing technology, specifically computers, in almost every facet of their lives. This researcher focused on student usage of computers for testing, in the education arena, particularly state-mandated EOC high-stakes tests.

During the course of a high school Sociology class, where I assigned my students a social research project for their final, it was brought to my attention that perhaps not all students were comfortable taking tests on computers, or that not all students score their best on computer-based tests (CBTs). Working with two students as part of their final, I helped them prepare a study using the scientific method to test a hypothesis concerning CBT versus paper-pencil (PPT). Interest in this topic occurred because of the, then recent, Missouri Department of Elementary and Secondary Education (MODESE) requirement that EOC tests be CBT. The research students selected a class and had students in the class take a test PPT and then CBT. Students were asked to include their cumulative grade point average (CGPA). Although this beginning attempt at social research had some errors in planning and documentation, their study did seem to show a

relationship between CGPA and mode of testing.

I was curious as to whether on a larger scale, and using statistical analysis, there could be a documented study showing a relationship between CGPA and mode of testing. Each time an EOC testing period occurred, all school computers were inaccessible for classroom use for a two week period. Usually this time period coincided with finals preparation. Myself as well as other faculty members questioned why the need to go to CBT. Allowing some students to take PPT would free up some of the computer time. Cost and grading were definitely factors in the state's decision, but I felt that student achievement, in the guise of EOC scores, may not have been considered to a great extent, when the decision was made to mandate CBT. If a relationship existed between CGPA and a student's EOC score based on mode of test administration, the researcher felt a study would yield valuable information. Holding to the tenets of differentiation, I also felt it would be informative to see how mode of testing affected student achievement, and how this might be transferred to classroom testing.

During the process of literature searching for Chapter 2, I was unable to find any study that contemplated the relationship of student CGPA and EOC test scores, or any study with regard to CGPA and test mode administration. Therefore, it was felt this dissertation could fill a research gap by providing material not previously studied.

Data Analysis

I used quantitative sources, student CGPA and EOC scores, in my research. Working with the hypotheses that students with a 2.5 or higher CGPA would score higher on PPT, while those with less than 2.5 CGPA would score higher on CBT, I sought to establish a correlation between CGPA and mode of test administration. Data was

requested from two large St. Louis suburban schools districts for the 2009-2010 school year. The 2009-2010 school year was selected for data gathering as it was the last school year that MODESE allowed districts a choice of test mode. Starting in the 2010-2011 school year, school districts in Missouri were required to administer test using CBT.

Research Questions

In this study I focused on several questions.

1. What is the relationship between students' cumulative grade point averages and their exam scores for United States Government End-of -Course exams, through paper-pencil or computer-based administration?
2. What is the relationship been students' cumulative grade point averages and their exam scores for United States Government End-of-Course exams, through paper-pencil or computer-based administration, disaggregated by No Child Left Behind subgroups?
3. Should the Missouri Department of Elementary and Secondary Education continue with their mandate for computer-based administration of End-of-Course exams, or would choice of mode of administration be in the best interest of students; better demonstrating student mastery of the curriculum as determined by their performance on these exams, and therefore benefit their district and the state?

Through Null Hypothesis 1 this study demonstrated there were no differences in average student test scores based on testing mode, when considering performance of the overall sampling of data, on the United States Government EOC exam. Exceptions were found for the female, Asian, White, and Hispanic subgroups. Data for each of these

indicated statistical differences in student averages. Should an additional study be done along these lines it would be worth seeing if these subgroups were still the exception. With data to support the Alternative Hypotheses 2 and 3, I found there was a relationship between students CGPA and EOC scores when taking tests PPT or CBT. Students with a higher GPA yielded a higher EOC score. The only exception was for the White subgroup participating in PPT administration of the EOC. This relationship between CGPA and EOC scores is to be expected as students tend to have similar scores across the board whether taking classroom or high-stakes tests, doing daily work, or calculating their CGPA.

Non-rejection of Null Hypothesis 4 for all subgroupings except males and students working with Individual Education Plans (IEP) resulted in a lack of support for Alternative Hypothesis 4, which analyzed whether or not students with CGPAs of 2.5 or above would score higher than those with a CGPA of less than 2.5 on PPTs. Additionally, non-rejection of Null Hypothesis 5 for all subgroupings except male, non-Free Reduced Lunch (FRL), and students working with an IEP resulted in lack of support for Alternative Hypothesis 5, which analyzed whether or not students with CGPAs below 2.5 will score higher than those with a CGPA of 2.5 or above on CBTs. Based on Hypotheses 4 and 5, I am able to state that the overall premise of this study, that students with CGPAs above 2.5 will do better taking PPTs, while students below 2.5 will do better taking CBTs, is invalidated, except for the two subgroups represented by males and students assigned an IEP.

One of the reasons students with IEPs validate Hypotheses 4 and 5 could be that they have the opportunity to take the EOCs under different test conditions than non-IEP

students. IEP students may have the test read to them and may also use additional testing time.

After this study I believe the answers to questions one and two are: students CGPAs will be directly related to their EOC scores, but their CGPAs will not be predictive of what mode of test administration would best demonstrate students' mastery of curriculum. Finding CGPAs directly related to EOC test scores is to be expected, as stated previously, students tend to have similar scores across the board whether taking classroom or high-stakes tests, doing daily work, or calculating their CGPA.

With regard to question three the study demonstrates that either CBT or PPT would be acceptable modes of administration for the EOC tests, and that other considerations may dictate the states need to mandate computer testing.

Implications and Recommendations for Further Study

While sharing information regarding my study with a colleague, we began discussing how this data would translate into classroom testing. He was so intrigued by the idea of a relationship between CGPA and test mode that he is in the process of conducting a small study with two of his U.S. History classes. He will be using two similar tests and have students take one test PPT and the other CBT to decide if this study's hypothesis holds true for classroom testing.

I too plan to test this hypothesis in the classroom. As a classroom teacher, I am particularly interested in trying this in the classroom because of the recent emphasis on differentiation. I believe not just differentiation in teaching methods, but differentiation in testing procedures, could make a change in classroom evaluation of student achievement, allowing focus to be centered on testing for content knowledge, rather than

test-taking skills.

In a recent Sociology class I taught, the students and I were discussing their semester self-directed social research finals, when the topic of my dissertation arose. One group, besides being interested in my hypothesis, became interested in whether students would rather take the EOC tests CBT or PPT. For their final they gave out surveys to U.S. Government classes asking whether they would rather take the EOC, PPT, or CBT. The overwhelming response was that they would like to take the test CBT. This result, in comparison to the results of my students, leads me to believe that students do not even understand their strengths regarding CBT or PPT. It seemed to me the CGPA would be an acceptable guideline for decision making regarding type of testing.

To this point I have only discussed high school EOC or classroom testing, but we need to look at a broader range of testing. For instance, in the high-stakes testing arena, tests like the ACT, SAT, or LSAT that are only offered paper-pencil may not measure a students' true level of educational maturity. Likewise, other high-stakes tests such as the MCAT, NICET, and FCAT, which are now taken by computer, may be a disadvantage for some

I recommend additional studies be conducted nationwide to continue to examine the correlation between CGPA and mode of testing. These additional studies should be conducted with EOC tests, classroom tests, and even high stakes tests such as the ACT. Since very little information was found in the literature on this topic, it bears another study to see if the results from this study will be upheld.

I would like to see an additional study using data from smaller, urban, and rural school districts to see if the results are similar to those from the two suburban school

districts used in this study. Students in the districts used in this study had early and recurring access to computers throughout their schooling and testing careers; which may not have been provided for students in smaller or less economically advantaged districts.

Another area for future study would be to develop a study using classroom testing; have students with CGPA below 2.5 take their tests on CBT, and those with CGPA of 2.5 and above take their tests PPT, for a semester, to see if their semester grades show an improvement over previous semester grades. Or, have all students take similar tests in the CBT and PPT formats and then analyze with their CGPA.

Besides testing in the classroom using students selected by their CGPA to decide what method of exam would be taken, it would be beneficial to do an experimental study placing students in EOC testing conditions by their CGPA.

As this study only examined CGPA from the 2.5 range (below and above), I think another study disaggregating CGPA into further categories would give a more specific indication as to at what point based on CGPA, at which moving to CBT or PPT would be a more accurate measure of a students' achievement.

Use of Technology

Working as a teacher for the last 12 years, I have observed a strong push towards the use of technology even when more traditional methods of teaching have been successful. We run thousands of copies off and hand them out to students, we provide websites, computers, Smartboards, airliners, Senteos, blogs and many other forms of technology. Students no longer write the notes or assignments off the board, or read books for research. There is a push for digital videos, Powerpoints and podcasts in place

of poster board and oral presentations. There is a place for technology, but it does not have to be at the expense of former teaching methods.

Then we have the classroom teachers who give scantron tests because they are quick to grade. In order to get a clearer picture of student achievement some type of performance assessment is helpful. Students often complain about the number of Powerpoints they are given as assignments and the number of Powerpoints they have to sit through in class given by both teachers and classmates. The dynamic teacher who used group work, demonstrations, maps, reenactments, and storytelling to keep students interested enough to learn seems to be fading from the educational scene. Educational institutions and teachers in leadership positions need to help education find a balance between traditional methods of teaching and the use of technology in the classroom.

Results from this study indicated for Hypothesis 1, no difference in average EOC scores between paper-pencil and computer usage. We should look at the reasons for this consistent push to do or use all things computer. Perhaps students would enjoy and learn more from storytelling time versus a YouTube video, or learn by reading an actual book rather than a computer version, or allowed to draw their own pictures on a dry-erase board rather than use a Smartboard application, or give a presentation with paper notes and physical objects rather than a Power point. We should encourage our teachers to use all the tools available to them and their students, rather than just those for computer based lessons.

Even teacher evaluation plans have a number of sections for grading how the teacher uses technology, which encourages more usage of computers in lesson planning (Parkway School District, 2012*)*. There are abundant professional development sessions

available dealing with the use of the computer in the classroom. Whether they are building, district, or conference opportunities, teachers are strongly encouraged to attend as many as possible. National conventions for various curriculum stress technology sessions.

School districts now use on-line grading systems for teachers, parents and students. In many cases these systems have done away with face to face parent teacher communication. Parents no longer feel the need to attend school conferences as they can track their students' grades on-line, see what work they are missing, days absent, and information concerning their behavior. Although this is a fast method of grading for teachers and a quick way for parents to track their students, we are definitely missing what we had in the past when teachers and parents spent some time discussing the students' needs. We have lost this human touch to our work as educators.

There are definitely some technological tools that can be useful in the classroom, such as the computer for writing a draft for an assigned paper, DVDs, and projection devices. But educators need to question whether they must have these tools and whether the monetary cost is equivalent to improved student learning. We should question whether every classroom needs a SMART board. There are many things we do on the SMART-board that can be done in another way. Much of our school budgets are centered on technology. A school can be thought less of because the district cannot afford technology and continues to use paper-pencil methods. Programs and services have gone by the wayside so that more technology could be integrated in schools. As teachers we need to question the current trend in education that indicates the ranking of a school based on the amount of technology available and the number of programs offered using

this technology ("STEM Schools Best High Schools," 201 2). I do not believe that computers are the answer to everything and feel that sometimes students will be more successful using paper and pencil. Out of curiosity I asked my regular freshmen students if they preferred testing on a computer or with paper-pencil. They answered computer. Asking my honors freshmen the same question, their answer was paper-pencil. When ask why they preferred paper-pencil they talked about preferring to spread out their papers and sources and being able to see everything at once.

There has been a strong shift in education toward the use of computers, but we need to encourage our teachers to use other methods of educating, including some from the past. College and university teacher education programs stress the need to use technology in the classroom. Perhaps they have jumped on the band wagon a little too forcefully and in the future need to back off and present students with other options for teaching methods.

Practitioner Applications

I would hope that any teacher reading this study would think about the type of testing they do in their classrooms and consider giving their students a choice of CBT or PPT, or at least giving both types of assessments depending on the curriculum. Part of my school districts teacher evaluation program examines delivering formative and summative assessments in multiple formats, as well as individualizing student feedback.

Adjusting the type of lessons taught in the classroom to include more than computer based lessons, could also be beneficial to those students who perform better with paper-pencil applications.

Conclusion

The randomly sampled objective data, and the correlational study based on this data, suggests that the hypothesis that students with a 2.5 or higher CGPA will score higher on EOC PPTs, while students with less than a 2.5 CGPA will score higher on EOC CBTs, is not supported for the overall sampling of data. Results differed for smaller subgroupings. This study determined that students with higher CGPAs will also have higher EOC scores.

Acknowledging NCLB's directive to states to raise student achievement levels, mode of test administration for standardized high-stakes tests such as Missouri's EOCs becomes relevant. Although findings in this study allude to no relationship between CGPA and mode of test administration; because this appears to be the first study of its kind, it is felt additional studies are merited. The possibility of transferring this prediction to classroom testing to further aid teachers in their application of differentiation in the classroom, thus improving student test scores and therefore CGPA, and provide overall a positive impact on student performance; is another reason for conducting additional studies on these premises.

References

ACT. (2010). Preparing for the ACT. Retrieved June 21, 2011, from http://www.act.org/aap/pdf/preparing.pdf

Al-Amri, S. (2008). Computer-based testing vs. paper-based testing: a comprehensive approach to examing the comparabilty of testing modes. *Essex Graduate Student Papers in Language & Linguistics, 1*0, 22-44.

Barker, L. (n.d.). Ehow.com. Retrieved April 30, 2011, from http://www.ehow.com/about_5063650_map-testing.html.

Barth, P., & Mitchell, R. (2006). Standardized tests and their impact on schooling: Q & A. *Center for Public Education* Retrieved September 7, 2010, from http://www.centerforpubliceducation.org/site/app/nlnet/content3.aspx?c=lvlXliNOJwE&b

Bartlett, J. E., Alexander, M. W., & Ouwenga, K. (2001). A comparison of online and traditional testing methods in an undergraduate business information technology course. OSRA. Retrieved March 3, 2010,from http://www.osra.org/2001/bartlett1.pdf

Baxter, J. H., Hungerford, B. C., & Helms, M.M. (2010). Predicting success in the introduction to computers course: GPA vs. student's efficacy score. *Information Systems Educators Conference Proceedings*, 1-22.

Bernt, F. M., Bugbee, A. C., & Arco, R. P. (1990). Factors influencing student resistance to computer administered testing. *Journal of Research on Computing in Education*, *22*(3), 265-275.

Bluman, A. G. (2008). *Elementary statistics:A brief version* (4th ed.). New York, NY:

McGraw Hill.

Bodmann, S. M., & Robinson, D. H. (2004). Speed and performance differences among computer-based and paper-pencil test [Electronic version]. *Journal of Educational Computing Research, 31* (1), 51-60.

Brown, G. (n.d.). *Early data reveal GPA as better predictor of college success*. NCAA. Retrieved May 3, 2010, from http://www.ncaa.org/wps/wem/connect/neaa/NCAA/NCAA+News/NCAA+News+Online/2

Carpenter, D. (2005, May 16). *GRE score and undergraduate GPA as predictors of success in graduate school*. Evangel University Business and Economics Department. Retrieved August 10, 2010, from http://www.docstoc.com/docs/49052345/GRE-and-UGPA-as-Predictors-of-Success-in-Grad.

Choi, S. W., & Tinkler, T. (2002, April). *Evaluating comparability of paper-and-pencil and computer-based assessment in a K-12 setting*. National Council on Measurement in Education. Retrieved April 8, 2010, from http://www.ncme.org/repository/incoming/100.pdf

Clariana, R. W., & Wallace, P. (2002, November). Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology*, *33*(5), 593-602.

Clipper, A. (2010, June 4). *GPA requirements for college admission*. Ehow.com. Retrieved May 22, 2011, from http://www.ehow.com/list_6626595_gpa-requirements-college-admission.html

College admission requirements and your GPA. (n.d.). *Peterson's College Search.* Retrieved May 22, 2011, from http://www.petersons.com/college-search/college-admission-requirements-gpa.aspx

College degree guide. (2010). *About the Graduate Record Examination.* Retrieved January 29, 2010, from http://www.collegedegreeguide.com/articles-fr/gre.htm

Computer-based assessment. (2010). Wikipedia. Retrieved February 1, 2010, from http://en.wikipedia.org/wiki/Computer-based_testing

The Computer-based or online administration of paper-and-pencil tests. (2003, April 11). CTB/McGraw-Hill. Retrieved August 15, 2011, from www.ctb.com/ctb.com/control/openFileShowAction?mediaId=666.0

Dahlman, A., Hoffman, P., & Brauhn, S. (n.d.). Classroom strategies and tools for Differentiating Instruction in the ESL Classroom. Retrieved December 12, 2011 from http://minnetesol.org/journal/vol25_html_pages/17_Dahlman.htm

DeBerard, M. S., Spielmans, G., & Julka, D. (2004, March). Predictors of academic achievement and retention among college freshmen: a longitudinal study. *College Student Journal*, 38(1), p, 66.

Defining adequate yearly progress (AYP). (n.d.). *North Central Regional Educational Laboratory:* Retrieved February 1, 2010, from http://www.ncrel.org/sdrs/areas/issues/content/cntareas/science/sc7ayp.htm

Desmarais, L., Woble-Valenski, M.A., & Oestmann, E. (2011). Factors influencing physical therapist assistant licensure examination success. *Journal of Physical Therapy Education*, *25*(2), 36-42.

Dillon, A., McKnight,C., & Richardson, J. (1988). Reading from paper versus reading from screen. *The Computer Journal* , *31*(5), 457-464.

Education.com. (n.d.). Remedial Education. Retrieved May 22, 2011, from http://www.education.com/definition/remedial-education

ETS GRE. (n.d.). *Here are the top frequently asked questions (and answers) about the GRE General Test.* Retrieved June 21, 2011, from http://www.takethegre.com/gre-general-test-top-faqs

ETS the Praxis Series. (2011). The Praxis for Test Takers. Retrieved June 21, 2011, from http://www.ets.org/praxis/register/

Ferguson Florissant School District. (2011, March 7). *Letter to parents.* Retrieved May 1, 2011, from http://www.fergflor.org/11792082016512807/lib/11792082016512807/EOC_Letter_to_Parents_3-4-11.pdf

Florida Department of Education. (n.d.). *Frequently asked questions*. Retrieved June 21, 2011, from http://www.fldoe.org/faq/default.asp?Dept=179&ID=1412#Q1412

Florida Department of Education. (2006, September 4). *What do we know about choosing to take a high-stakes test on a computer*? Retrieved May 15, 2010, from http://www.fldoe.org/asp/k12memo/pdf/WhatDoWeKnowAboutChoosingToTakeAHighStakesTestOnAComputer.pdf

Gallagher, A., Bridgeman, B., & Cahalan, C. (2002). The effect of computer-based tests on racial-ethnic and gender groups. *Journal of Educational Measurement* [Electronic file] , *39*(2), 133-147.

Geiser, S., & Santelices, M. (2007). *Validity of high-school grades in predicting student success beyond the freshman year.* Retrieved March 8, 2010, from http://cshe.berkeley.edu/publications/docs/rops.geiser._SAT-6.12.07pdf

Goertz, M. E. (n.d.). *The federal role in defining "adequate yearly progress*:" *The flexibility/accountability trade-off*. Retrieved July 30, 2011, from http://www.cpre.org/images/stories/cpre_pdfs/cep01.pdf

Grade (education). (2010). Wikipedia. Retrieved February 1, 2010, from http://en.wikipedia.org/wiki/Grade_(education)

Gray, L. (2010, April 10). Debate. *New York Times Upfront Magazin*e. *142*(13), p. 30.

Great Schools. (n.d.a). *Testing in Missouri: an overview*. Retrieved July 22, 2011, from http://www.greatschools.org/students/local-facts-resources/419-testing-in-MO.gs?page=1

Great Schools. (n.d.b). *What the No Child Left Behind Law Means for Your Child*. Retrieved August 7, 2011, from http://www.greatschools.org/improvement/quality-teaching/61-no-child-left-behind.gs

Hall, T., Strangman, N., & Meyer, A. (2003). *Differentiated instruction and implications for UDL implementation*. National Center on Accessing the General Curriculum. Retrieved February 25, 2010, from http://aim.cast.org/sites/aim.cast.org/files/DI_UDL.1.14.11.pdf

Heubert, J. P., & Hauser, R. M., (ed.) (1999). *High stakes testing for tracking, promotion, and graduation.* National Academy Press. Retrieved June 19, 2009, from http://www.nap.edu/openbook.php?record_id=6336&page=2

How to calculate your grade point average (GPA). (n.d.). Back to College. Retrieved September 12, 2010, from http://www.back2college.com/gpa.htm

I feel like a number: gradepoint average and admissions. (2000, September 1). Retrieved September 7, 2010, from https://pantherfile.uwm.edu/ccp2/www/work/gpa.html

Insight Assessment. (n.d.). Paper & pencil testing. Retrieved June 26, 2011, from http://www.insightassessment.com/Services/Paper-and-Pencil-or-Online-Testing/Paper-Pencil-Testing

International Student Guide to the United States of America. (n.d.). TOEFL, GMAT, GRE, LSAT ,MCAT, SAT, ACT... Which tests do you need? Retrieved January 29, 2010, from http://internationalstudentguidetotheusa.com/articles/which_tests.php

Jackson School District. (n.d.). Why Co-Teach? Retrieved May 22, 2011, from http://www.up140.jacksn.k12.il.us:8080/Smith//readingpix.htm

Kapes, J. T., Martinez, L., Chui-Fung, I., Slivinski, T., & Hardwick, S. (1998) Internet-based vs. paper-pencil occupational competency test administration: an equivalency study. *Journal of Vocational Education Research* [Electronic file] , *23*(3), 201-219.

Kaplan Test Prep. (n.d.). What is Changing on the MCAT? Retrieved May 28, 2011, from http://www.kaptest.com/oneoff/mcat-test-change/computer-based-mcat-information.jhtml

Kim, D., & Huynh, D. (2007). Comparability of computer and paper-and-pencil versions of Algebra and Biology assessments (Electronic Source). *Journal of Technology, Learning, and Assessment*, 1-30.

Kingore, B. P. (2005). *Differentiating instruction: Rethinking traditional practices.* Retrieved May 28, 2011, from http://www.bertiekingore.com/diffinstruct.htm

Leeson, H. (2006). The mode effect: a literature review of human technological issues in computerized testing. *International Journal of Testing, 6*(1), 1-

Maguire, K. A., Smith, D.A., Brallier, S. A., & Palm, L. J. (2010). Computer-based testing: a comparison of computer-based and paper-pencil assessment. *Academy of Educational Leadership Journal*. 14(4), 117-125.

Mayes, D. K., Sims, V K., & Koonce, J. M. (2001). Comprehension and workload differences for VDT and paper-based reading. *International Journal of Industrial Ergonomics, 28*(6), 367-378.

McClarty, K. L. & Davis, L. L. (2006). *Secondary analysis method s in compoarability research: A review of methods used with the Texas assessment of knowledge and skills.* Pearson Educational Measurement. Retrieved February 1, 2010, from http://www.pearsonedmeasurement.com/donwloads/conference/SecAnalysisCompRes_cp0602.pdf

McGraw Hill. (2009, December). *Missouri assessment program grade-level assessments.* MODESE. Retrieved April 30, 2011, from http://www.dese.mo.gov/divimprove/assess/tech/documents/2009-MAP-Technical-Report.pdf

Meador, D. (n.d.). *High stakes testing*. About.com Teaching. Retrieved May 8, 2011, from http://teaching.about.com/od/gloss/g/High-Stakes-Testing.htm

Millsap, C. (2000). *Comparison of computer testing versus traditional paper and pencil testing*. Retrieved August 8, 2010, from

http://www.library.unt.edu/theses/open/20002/millsap_claudette-m/dissertation.pdf

Missouri Department of Elementary and Secondary Education. (n.d.). Social Studies Grade- and Course-Level Expectations 2.0. Retrieved March 25, 2011, from http://dese.mo.gov/divimprove/curriculum/GLE/documents/ss_gle_2.0_k8_0907.pdf

Missouri Department of Elementary and Secondary Education. (2007, January). Graduate requirements for students in Missouri public schools, effective for graduates of the class of 2010. Retrieved May 22, 2011, from http://dese.mo.gov/divimprove/sia/Graduation_Handbook_2010.pdf

Missouri Department of Elementary and Secondary Education. (2008a, October 17). *Curriculum frameworks - preface*. Retrieved August 8, 2010, from http://dese..mo.gov/divimprove/curriculum/frameworks/preface.html

Missouri Department of Elementary and Secondary Education. (2008b, April 25*). End-of-course exams frequently asked questions*. Retrieved August 5, 2010, from http://dese.mo.gov/divimprove/assess/documents/EOCFrequentlyAskedQuestions.pdf

Missouri Department of Elementary and Secondary Education, (2008c, October 31). *Overview of knowledge standards*. Retrieved August 8, 2010, http://desemo.gov/standards/content.html

Missouri Department of Elementary and Secondary Education. (2008d, October 31). Show-Me Standards. Retrieved August 8, 2010, from http://dese.mo.gov/standards/content.html

Missouri Department of Elementary and Secondary Education, (2009a) *End-of-course assessment, guide to interpreting results*. Retrieved January 15, 2010, http://docs.google.com/viewer?a=v&q=cache:Hv9dn2O7V5sJ:dese.mo.gov/divimprove/assess/documents/EOC_2010_GIR.pdf+missouri+end-of-course+assessment+gide+to+interpreting+results&hl=en&gl=us&pid=bl&srcid=ADGEEShqbBYJveDdYaRTRD1A-mJFpQECB_9obivEIgAJ1708vYrCHgr

Missouri Department of Elementary and Secondary Education. (2009b*). End-of-course assessment, online test examiner's manual*. Retrieved February 15, 2010, from http://docs.google.com/viewer?a=v&q=cache:W2_8-UJGB2sJ:web.sjsd.k12.mo.us/Assessment/documents/StJosephEOCFall2009Oct272009_2.ppt+missouri+end-of-course+assessment+online+test+examiner's+manual&hl=en&gl=us&pid=bl&srcid=ADGEESgMg0KiSiLGNI_x1yyYoWLZN1FW3kGS

Missouri Department of Elementary and Secondary Education. (2009c). *End-of-course assessment test coordinator's manual training, phase I and phase II 2009/2010*. Retrieved May 8, 2011, from http://map.missouristate.edu/assets/MAP/EOC_Manuals_Training_2009-2010.pdf

Missouri Department of Elementary and Secondary Education, (2009d). *Missouri end-*

of-course assessment. Retrieved May 29, 2011, from http://dese.mo.gov/divimprove/assess/documents/2009-2010_EOC_Online_TEM_000.pdf

Missouri Department of Elementary and Secondary Education. (2009e). *Missouri end-of-course assessments technical report phase 1 assessments 2008-2009*. Retrieved February 15, 2010, from http://dese.mo.gov/divimprove/assess/tech/documents/2008-2009-EOC-TechReport.pdf

Missouri Department of Elementary and Secondary Education (2009f) *Missouri end-of-course paper/pencil versus online comparability study*. Jefferson City, Missouri: Riverside Publishing.

Missouri Department of Elementary and Secondary Education. (2009g, August 12). *State Officials Pleased with Steady Gains on MAP Tests*. Retrieved May 6, 2011, from http://dese.mo.gov/news/2009/mapeoc.htm

Missouri Department of Elementary and Secondary Education. (2010a). *Guide to interpreting results*. Missouri End of Course Assessments Missouri Department of Elementary and Secondary Education Guide to Interpreting Results. Retrieved January 17, 2011, from http://www.dese.mo.gov/divimprove/assess/documents/EOC_2010_GIR.pdf

Missouri Department of Elementary and Secondary Education. (2010b*) Missouri end of course assessments technical report phase II assessments 2009-2010*. Retrieved August 22, 2011, from http://dese.mo.gov/divimprove/assess/tac/documents/asmt-eoc-p2-tech-report-2010.pdf

Missouri Department of Elementary and Secondary Education. (2010c, February 11). *Show-me standards*. Retrieved March 14, 2011, from http://dese.mo.gov/standards/process.html

Missouri Department of Elementary and Secondary Education. (2010d, July 23). *Understanding your annual yearly progress (AYP) report 2010-2011*. Retrieved August 8, 2010, from http://dese.mo.gov/divimprove/sia/dar/UnderstandingYourAYP.pdf

Missouri Department of Elementary and Secondary Education. (2011a, January 13). *About office of college and career readiness- curriculum/staff*. Retrieved February 22, 2011, from http://dese.mo.gov/divimprove/curriculum/aboutus.htm

Missouri Department of Elementary and Secondary Education. (2011b, January 13). *About the Missouri assessment program*. Retrieved February 22, 2011, from http://dese.mo.gov/divimprove/assess/staff.html

Monroe County Schools. (2001, September 4). *Graduation requirements*. Retrieved May 22, 2011, from http://www.monroecountyschoolswv.org/policies.php?pd_id=21

Montgomery, M. (2009, May 27). *Grades your GPA, education, and learning: how do we compare apples to apples*? Retrieved February 20, 2010, from http://greatcollegeadvice.com/grades-your-gpa-education-and-learning-how-do-we-compare-apples-to-apples/

Morin, A. (n.d.). *What is high stakes testing*? About.com School-Age Children. Retrieved May 8, 2011, from http://childparenting.about.com/od/schoollearning/a/high-stakes-tests-definition.htm

National Center for Fair and OpenTesting. (2007, August 20). *Computerized testing: More questions than answers.* Retrieved August 10, 2010, from http://fairtest.org/computerized-testing-more-questions-answers

National Institute for Certification in Engineering Technologies. (2008, March). NICET is Transitioning to Computer Based Testing. Retrieved May 22, 2011, from http://www.nicet.org/newsletter/march2008/computer_basedtesting.cfm

Nejad, M. A. (1995). *Affects of age and GPA on learning electronics via computer simulation-based and traditional instruction*. Annual Conference of the Mid-South Educational Research Association.

Nichols, S. L., & Berliner, D.C. (2008, May). Why has high-stakes testing so easily slipped into contemporary American life? *Phi Delta Kappan*, *89*(9), pp. 672-676.

Northwestern University, School of Education and Social Policy. (2010). Center for Talent Development. Retrieved July 22, 2011, from http://www.ctd.northwestern.edu/resources/displayArticle/?id=134&pf=0

OneStopGRE. (2009, November 11). GRE: Computer-based test vs paper-pencil-based test. Retrieved May 22, 2011, from http://forum.onstopgre.com/forum_posts.asp?TID=445

Paek, P. (2005). *Recent trends in comparability studies*. Pearson Educational Measurement. Retrieved February 1, 2010, from http://wwwstage.pearsonedmeasurement.com/donwloads/researc/TrendsCmpStdies_rr0505.pdf

Parkway School District. (2010). This is Parkway. Retrieved April 8, 2010, from http://www.this-is-parkway.com/2010/default.asp

Parshall, C. G., & Kromrey, J. D. (1993). *Computer testing versus paper-and-pencil testing: an analysis of examinee characteristics associated with mode effect.* Retrieved from ERIC.

Pearlman, M. (2001). High-stakes testing: perils & opportunities, remarks presented at the Pennsylvania Education Policy Forum, Harrisurg, PA. Retrieved February 25, 2010, from http://www.eplc.org/mpearlman.html

Poggio, J., Glasnapp, D. R., Yang, X., & Poggio, A. J. (2005). A Comparative Evaluation of Score Results from Computerized and Paper & Pencil Mathematics Testing in a Large Scale State Assessment Program [Electronic version]. *The Journal of Technology, Learning, and Assessment,* 3(6), 1-30.

Pommerich, M. (2004). Developing computerized versions of paper-and-pencil tests: mode effects for passage-based tests [Electronic version]. *The Journal of Technology, Learning, and Assessment*, *2*(6), 1-44.

Public Law 107-110. (2002, Jan. 8). Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

Puhan, G., Boughton, K., & Kim, S. (2007). Examing differences in examinee performance in paper and pencil and computerized testing [Electonic version]. *The Journal of Technology, Learning and Assessment*, *6*(3), 1-20.

Rockwood School District. (2010). Rockwood demographics. Retrieved March 13, 2010, from http://ww.rockwood.k12.mo.us/aboutus/Pages/RockwoodDemographics.aspx

Russell, M., & Plati,P. (2000). *Effects of computer versus paper administration of a state-mandated writing assessment*. Chestnut Hill, ME.: Intasc

Schaeffer, G. A., Bridgeman, B., Golub-Smith, M. L., Lewis, C., Potenza, M. T., & Steffen, M. (1998). Comparabiliy of paper-and-pencil and computer adaptive test scores on theGRE General Test. Princeton, NJ: Educational Testing Service.

Sentance, M. (2000, Winter). Why do we need high-stakes assessments? *State Education Leader*, pp. 13-14.

Shead, M. (2006). *Using different parts of your brain*. Productivity 501. Retrieved September 20, 2010, from http://www.productivity501.com/using-different-parts-of-your-brain/88/

Sim, G., & Horton, M. (2005). *Performance and attitude of children in computer based versus paper based testing*. Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications (EDMEDIA) (pp. 3610-3614). Chesapeake, VA: AACE.

Sireci, S. G. (2009). *Standardized Testing is Useful*. Opposing Viewpoints Resource Center. Retrieved June 19, 2009, from http://find.galegroup.com/ovrc/retrieve.do?subjectParam=Locale%2528en%252C%252C...

Sloane, F., C., & Kelly, A.E. (2003). Issues in High-Stakes Testing Program. *Theory Into Practice, 42*(1), 12-17.

Slocum, B. (2009). *Computer based testing: the way of the future? Developing literacy in MST*. Retrieved April 14, 2010, from http://mstliteracy.wordpress.com/2009/10/08/computer-based-testing-the-way-of-the-future/=0&_userid=10&md5=1146bbb1206a0f2526cae672995e32a8

Stecher, B. M. (n.d.). Consequences of large-scale, high-stakes testing on school and classroom practice. In L. S. Hamilton, Making sense of test-based accountability in education (pp. 79-100). Santa Monica, CA: Rand.

STEM Schools Best High Schools. (2012). Retrieved August 20, 2012 from U.S. News and World Report: http://www.usnews.com/education/best-high-schools/national-ranking /stem

Teacher certification exams: what are the predictors of success? (2005, December 1). Goliath Business News. Retrieved September 7, 2010, from http://goliath.ecnext.com/coms2/gi_0199-5288778/Teacher-certification-exams-what-are.html

Tech's answer to testing; schools turn to computerized exams to meet new demands [Electronic version]. (2003, May 8). *Education Week*, *22*(35).

Texas Education Agency. (2008). *A review of literature on the comparability of scores obtained from examinees on computer-based and paper-based tests*. Retrieved August 10, 2010, from 214.97.204/WorkArea/linkit.aspx? Series: 198LinkIdentifier=id&ItemID

TOEFL. (2011, July 9). TOEFL step-by-step. Retrieved July 9, 2011, from http://www.toeflgoanywhere.org/content/all-about-toefl-test-what-you-need-know

Tran, K. (2009, April 28). *Students, faculty react to new end-of-course assessment exams*. Retrieved May 22, 2011, from http://www.columbiamissourian.com/stories/2009/9428/students-react-new-end-course-tests/

Troseth, K. (2008, September 29). *What colleges look for in applicants.* Suite 101 Retrieved May 22, 2011, from http://www.suite101.com/content/what-colleges-look-for-in-applicants-a70996

U.S. Department of Education. (n.d.). *Overview: Introduction - Why NCLB is important*. Retrieved July 19, 2011, from http://www2.ed.gov/print/elsec/leg/esea02/pg1.html

U.S. Department of Education. (2002, January 8). *Public Law 107-110.* Retrieved November 21, 2009, from http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

U.S. Department of Education. (2003). *No Child Left Behind A Parents Guide.* Retrieved August 12, 2012, from http://www2.ed.gov/parents/academic/involve/nclbguide/parentsguide.pdf

U.S. Department of Education. (2004). *NCLB executive overview*. Retrieved May 2, 2010, from http://www2.ed.gov/nclb/overview/intro/execsumm.html

U.S. Department of Education. (2009). *No Child Left Behind Standards and Assessments Peer Review Guidance*. Washington D.C., U.S. Department of Education.

University of South Florida, College of Education. (2009). Department of Educational Measurement and Research. Retrieved September 21, 2012 from http://www.coedu.usf.edu/main/departments/me/MeasurementandResearchStatisticalSoftwarePackages.html

Walker, S. F. (2000, Winter). High-stakes testing: too much? too soon? *State Education Leader*, pp. 1-3.

Wallace, P., & Clariana, R. B. (2005). Gender differences in computer-administered versus paper-based tests [Electronic version]. *International Journal of Instructional Media*, *32*(2).

Wan, L., Keng, L., McClarty, K., & Davis, L. (2009, December). *Methods of comparability studies for computerizezd and paper-based tests*. Pearson Test, Measurement & Research Services Bulletin Issue 10. Retrieved September 18, 2011 from http://www.pearsonassessments.com/NR/rdonlyres/7955109E-6724-4EC8-A25F-375FBE360D5F/0/Bulletin_10.pdf.

Wang, H., & Shin, C.D. (2009, November). *Computer-based & paper-pencil test comparability studies*. Pearson Test, Measurement & Research Services Bulletin Issue 9. Retrieved September 18, 2011 from http://www.pearsonassessments.com/NR/rdonlyres/93727FC9-96D3-4EA5-B807-5153EF17C431/0/Bulletin_9.pdf

Warrick, S. S., & Crumrine, R. S. (1986). Predictors of success in an anesthesiology residency. *Journal of Medical Education*, *61*(7), 591-595.

Way, W. (2006). *Online testing research informing and guiding transitions to computerized assessment*. Pearson Educational Measurement. Retrieved February 1, 2010, from https://pearsonassessments.com/NR/rdonlyres/E1A43E5F-D0C7-4CD7-93BE- EFFD863E9C2A/0/OnlineTestResearch_wp0601.pdf

Wenning, R. H. (2003, August). *No child left behind: Testing, reporting, and accountability*. Eric. Retrieved January 3, 2011, from http://www.ericdigests.org/2004-2/behind.html

Westchester Institute for Human Services Research. (2003, April). *High-stakes testing*. Retrieved from http://www.sharingsuccess.org/code/bv/testing.pdf

Williams, C. (2009, January). *Co-teaching and different styles of co-teaching*. Bright Hub. Retrieved March 25, 2011, from http://www.brighthub.com/education/special/articles/4096.aspx

Wise, S. L., Kingsbury, G. G., Thomason, J., & Kong, X. (2004, April). *An investigation of motivation filtering in a statewide achievement testing program*. National Council on Measurement in Education. Retrieved February 1, 2010, from http://www.jmu.edu/assessment/wm_library/Wise_K12_motivation_filtering_(1).pdf

Appendix A

PARKWAY SCHOOL DISTRICT
Application to Perform Research

Division of Program Evaluation
12860 Fee Fee Road
St. Louis, Missouri 63146
(314-415-5625)

I. Name of Primary Investigator: Barbara A. Ryan

Position: Teacher/Dept. Chair Affiliation: Parkway South High School
Office Address: 801 Hanna Rd., Manchester, Mo. 63021

Home Address: 113 Holly Green Dr., Ballwin, Mo. 63021

Office Phone: 314-415-5702 Home Phone: 636-394-6617

Names of additional members of research team: N/A

Name _____ Phone _____

Name _____ Phone _____

Name _____ Phone _____

II. Project Title
A Study of Student GPAs as predictors of success on United States Government End of Course Exams when using computer-based versus paper-pencil forms of administration

Description:
There are two purposes to this study. One is to determine if GPAs are a predictor of student success on computer-based and paper-pencil exams. The second purpose of the study is to possibly provide information to the state and school districts as to the benefits of one type of administration over the other; or the benefits for providing both options. The data to be used in this study will be the result of the end of course Government exams required for all students in United States Government courses in the state of Missouri. End of course exams are required by the state of Missouri in Math, English, Science and Social Studies. According to the Missouri State Board of Education the following are the purposes for the Missouri End-of-Course (EOC) Assessments:

* Measuring and reflecting student's mastery toward post-secondary readiness
* Identifying students' strengths and weaknesses

* Communicating expectations for all students
* Serving as the basis for state and national accountability plans
* Evaluating programs

Numerous computer-based versus paper-pencil testing comparability studies are available; but very few, if any, have evaluated the connections between GPA and type of testing. From a small project completed by students in their high school Sociology class, indications were that students with higher GPAs scored higher taking paper-pencil tests; while students with lower GPAs scored higher on computer-based tests.

A second purpose to the study is to provide information as to whether the Department of Secondary and Elementary Education's mandate, that starting in 2010-2011 school year all end of course testing will be done on computers, is beneficial to students and school districts. Information gained from the study may provide valuable information for the state in determining whether their mandate for computer-based administration should continue or are alternative methods of administration in the best interests of students and their performance on end of course exams. Finding computers for all students will be a strain for some districts.

Note: Please attach copies of any measures to be used (e.g. tests, questionnaires, surveys, etc.) N/A

III. Participant Involvement

Number of Subjects Time Requirements N/A

Pupils: _____ _____

Teachers: _____ _____

Administrators: _____ _____

Parents: _____ _____

Describe the involvement required of subjects (or access to records if subjects are not required).
Requesting all December 2009 and April 2010 U.S. Government end of course exam scores and correlating student GPAs broken down by No Child Left Behind Subgroups.

Number of persons visiting sites in connection with project: N/A

Frequency of visits during a school year: N/A

Total contact hours of the project: N/A

IV. Project Requirements

Number and type of school:

Elementary (K-5) _____ Middle school (6-8) _____ High school (9-12) __5___

Early Childhood Education (birth to kindergarten) _____

Adult Basic Education _____ Other _____ Grades required _____

Total number of schools __5_ Total number of classrooms _____

Other school characteristics:

_____

_____

Do you require any specific schools? Yes If yes, please provide names:

South High, West High, Central High, Fern Ridge, North High

Starting date of research: December 2009
Ending date of research: April 2010

Frequency of contact with subject: N/A

V. Results

What is the anticipated value of the research?

In general:
These assessments are intended to reflect a student's mastery and determine strengths and weaknesses; therefore it is important to discover which method of testing, paper-pencil or computer-based would be most beneficial to each individual. Since the end of course testing will be part of accountability, under No Child Left Behind, and serve as an evaluation of curriculum; students should take the test in the manner that will provide evidence of true mastery of the content. This will serve not only the student, but the district as well. If students with certain GPAs do better on computer-based or paper-pencil testing, should they then be tested in that manner?

To the Parkway School District:
As Parkway School District is held accountable under No Child Left Behind and is required by the state of Missouri to administer end of course exams; it is

beneficial to understand under what test conditions their students will best demonstrate mastery of the subject material. In addition these results could be applied to classroom assessments.

VI. Dissemination

How will the results of your study be used? Will they be available to the public in any form? If so, what groups will have access to the results? Will the Parkway School District, or any individuals within Parkway, be identified in your reports?

The results of my study will be published as part of my dissertation. In addition, the results will be made available to the Parkway and Rockwood school districts. No individuals will be identified in my report. Parkway will be mentioned as a data source.

VII. References (You may omit names, if you have promised confidentiality.)

Are other school systems involved in this research? Yes
Please name: Rockwood School District

Have you conducted research in other school systems? No

Please name

_____

Date(s)

_____

VIII. Human Subjects' Protection

Has this research been approved by a university or other institutional review for protection of human subjects?

Yes _____ No __X___

If yes, please indicate who reviewed the proposal and when:

_____

_____

If no, please explain why this proposal has not been reviewed for protection of human subjects: only requesting data, no access to human subjects

Note: All researchers who plan to collect information from or about individual students should attach copies of their proposed consent forms and a brief description of planned procedures for obtaining informed consent. Research involving individual students may

require the informed consent and signed agreement of parents or legal guardians.

4

IX. Upon completion of the research you will be required to submit two copies of the report (or summary) to the Director of Program Evaluation.

By signing this application, the applicant certifies that the research herein described involves an investigation which:

1. Promises to produce information of value to Parkway or the field of education;
2. Provides adequate safeguards for participants' rights;
3. Does not detract from the primary mission of instruction; and
4. Is not-for-profit in nature.

The documents can be expected by (date) May 2011

1. Barbara A Ryan — 3/30/10
signature of applicant — date

2. Lindenwood University — Dr. Vicki Hedges-Oldani
PRINT -- name of institution — institutional advisor, professor or supervisor

3. Vicki Hedges-Oldani — 636-946-2824
signature of advisor, professor or supervisor — office telephone

(For district use only)

1. _____ — 4/9/2010
signature of Director of Program Evaluation — date

2. _____ — _____
signature(s) of administrator(s) affected — date

extresapp(rev.3/05)
5

PARKWAY SCHOOL DISTRICT

Appendix B

ROCKWOOD SCHOOL DISTRICT

We will do whatever it takes to ensure all students realize their potential.

Administrative Annex
500 North Central
Eureka, MO 63025-1203

(636) 938-2215
(636) 938-2347 FAX
grahamerik@rockwood.k12.mo.us

Erik Graham
Director of Data Analysis and
Quality Management

M E M O R A N D U M—Research Request Approval

April 10, 2010

Barbara Ryan
Parkway South High School
801 Hanna Road
Manchester, MO 63021

RE: Request for Research within the Rockwood School District
Project Title: "A Study of Student GPAs as predictors of success on United States Government End of Course Exams when using computer-based versus paper-pencil forms of administration."
Institutional Affiliation: Lindenwood University

Ms. Ryan:

I have reviewed your application to perform research and the associated documentation in support of your application. Your application and associated documentation have met Rockwood School District guidelines for performing research and as such approval is granted to conduct the research project utilizing United States Government End of Course data and student GPA information derived from Rockwood School District data.

Your research proposal and accompanying documentation demonstrate you are aware of student, staff and school confidentiality issues and that you have taken precautions to protect student/staff/school privacy. Of specific note is that no individual student names will be shared with Lindenwood University. You have also indicated there will be no interference with the normal instructional time offered to students to conduct this research. You have also asked that the district provide the requested United States Government End of Course data and student GPA information to you. The district will provide this data to you in a format that protects individual student confidentiality and will meet your research needs. Thank you for being so thorough in outlining your procedures and your close attention to process details regarding the research.

Please be sure to contact Erik Graham, Director of Data Quality and Continuous Improvement sometime in August regarding the availability of this data and to clarify specific information requested regarding format and structure of the data files.

Lastly, upon completion of the research project, we look forward to the sharing of the results and findings of the research with our district. We are hopeful the research will be helpful in illuminating areas for improvement for our staff and students. If I can be of further assistance, please let me know. Good luck in your research investigation.

Sincerely,
Erik Graham

Director of Data Analysis and Quality Management

c: Dr. Carrie Luttrell, Executive Director of Curriculum, Instruction and Assessment

Appendix C

ROCKWOOD SCHOOL DISTRICT
Growing Together, Learning for Life

Administrative Annex
500 North Central
Eureka, MO 63025-1203
(636) 938-2215
(636) 938-2347 FAX
grahamerik@rockwood.k12.mo.us

Erik Graham
Director of Data Analysis and
Quality Management

APPLICATION TO PERFORM RESEARCH

Date: February 25, 2010

I. Name of Investigator: Barbara A. Ryan

Institutional Affiliation: Lindenwood University

Office Address: My Parkway School District address is: 801 Hanna Rd., Manchester, Mo. 63021

Home Address: 113 Holly Green Dr., Ballwin , Mo. 63021

Office Phone: 314-415-5702 Home Phone: 636-394-6617

If your research will be in partial fulfillment of a degree requirement, what degree is sought?
EdD in Instructional Leadership

II. Project Title: A Study of Student GPAs as predictors of success on United States Government End of Course Exams when using computer-based versus paper-pencil forms of administration

Description: There are two purposes to this study. One is to determine if GPAs are a predictor of student success on computer-based and paper-pencil exams. The second purpose of the study is to possibly provide information to the state and school districts as to the benefits of one type of administration over the other; or the benefits for providing both options. The data to be used in this study will be the result of the end of course Government exams required for all students in United States Government courses in the state of Missouri.

Proposed Starting Date: May 2010 – or as soon as data is available for the 2009/2010 school year

Proposed Completion Date: December 2010

III. Participants

Number of Subjects Required	Form of Participation (include treatments, tests, observations, etc.)
Students:	all 2009-2010 U.S. Government end of course exam scores for all Rockwood high schools, correlated with students GPA. No student names required. Also requesting data be broken down by NCLB subgroups.
Administrators:	
Teachers:	
Parents:	

Student Time Required: _____ per class/ per student

Time required of Other Participants: _____

E:\Rockwoods Data Request.doc

APPLICATION TO PERFORM RESEARCH – page two

If particular schools are being requested, please list: All Rockwood high schools

Cite one or two major studies that have been published or reported in your chosen area of research
At this time no studies have been located that directly relate GPA to end of course exams administered paper-pencil or computer-based.
Only one citation was found that discusses GPA and online versus paper-pencil testing.
Bartlett II, J., Alexander, M., & Ouwenga, Karen. *A comparison of online and traditional testing methods in an undergraduate business information technology course*. Retrieved February 12, 2010, from www.osra.org/2001/bartlett1.pdf

Attach a brief proposal of your research study explaining the important methodological features of the study (e.g., sampling methods, assessment tools, how confidential information will be handled, data analysis procedures, etc.). If you will be using a survey instrument, please enclose a complete copy.

IV. Results

What is the anticipated value of this research?

These assessments are intended to reflect a student's mastery and determine their strengths and weaknesses; therefore it is important to discover which method of testing, paper-pencil or computer-based would be most beneficial to each individual. Since the end of course testing will be part of accountability, under No Child Left Behind, and serve as an evaluation of curriculum; students should take the test in the manner that will provide evidence of true mastery of the content so they will be evaluated in a manner that best serves their learning style. If students with certain GPAs do better on computer-based or paper-pencil testing, should they then be tested in that manner?

V. IF YOU HAVE EXTENDED CONFIDENTIALITY, NAMES MAY BE OMITTED FROM THIS ITEM.

Have you conducted research in other school systems?
N/A

Please name: _____

Are other school systems involved in this research? Yes

Please name: Parkway School District

VI. UPON COMPLETION OF THE RESEARCH, YOU WILL BE REQUIRED TO SUBMIT TWO COPIES OF THE REPORT (OR SUMMARY).

A MEMORANDUM INDICATING PROCEDURAL PROBLEMS, UNUSUAL EXPERIENCES, RECOMMENDATION, COMMENTS AND OBSERVATIONS WOULD ALSO BE WELCOMED.

The documents can be expected by (date): May 2011

1) Barbara A Ryan — 3/30/10
(Signature of Applicant) (Date)

2) Dr. Vicki Hedges Oldani — Lindenwood University
(PRINT name of institutional advisor, Professor or supervisor) (Institution)

3) _____ — 636-946-2821
(Signature of advisor, professor or supervisor) (Office Telephone)

C:\Users\Vicki\AppData\Local\Temp\Rockwoods Data Request-1.doc

Brief proposal of research study as requested above:
Methodology/procedures:

* Application will be made to Parkway and Rockwood School Districts to obtain data.
* The data requested will be all United States Government end of course exam scores for the 2009/2010 school year.
* For each exam score the student's matching GPA will be requested.
* No student names will be required.
* The data will be requested broken down by No Child Left Behind subgroups.
* Quantitative statistical analysis will be done to test the hypothesis.
* Data will be stored on computer until the completion of the dissertation at which time it will be deleted

Vitae

Barbara Alice (Binder) Ryan earned her Bachelor of Arts degree from the University of Missouri-St. Louis, with a major in Sociology and a minor in Psychology, in 1970. After working for Ralston Purina Company for 10 years, her last role as Research Librarian, Mrs. Ryan took some time off to be a stay-at-home mom. She then returned to the workforce in the Parkway School District in various capacities including principal's secretary, substitute teacher, and teaching assistant for 10 years, during which time she returned to school to earn her teaching certification. She completed coursework toward teacher certification at the University of Missouri – St. Louis in 2000.

For the following 12 years, she worked as a classroom teacher at Parkway South High School and for four of those years as a Social Studies department chair. She has presented at state and national Social Studies conferences, served on many district committees, and sponsored numerous student organizations including Senior Class, Key Club, Special Olympics, Renaissance, and Youth In Government. During this time, Mrs. Ryan earned her Master of Arts in Educational Administration in 2006 from Lindenwood University and her anticipated graduation date from Lindenwood University's Doctoral Program in Instructional Leadership is May, 2013.