END OF COURSE GRADES AND END OF COURSE TESTS IN THE VIRTUAL ENVIRONMENT: A STUDY OF CORRELATION

By

Jamie Gilbert Philipp

Liberty University

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University

March, 2014

END OF COURSE GRADES AND END OF COURSE TESTS IN THE VIRTUAL ENVIRONMENT: A STUDY OF CORRELATION

By: Jamie Gilbert Philipp

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University, Lynchburg, VA

March, 2014

APPROVED BY:

Vivian O. Jones, PhD, Committee Chair

Marlynn M. Griffin, PhD, Committee Member

Steven McDonald, MBA, EdD, Committee Member

Scott B. Watson, PhD, Associate Dean, Advanced Programs

END OF COURSE GRADES AND END OF COURSE TESTS IN THE VIRTUAL ENVIRONMENT: A STUDY OF CORRELATION

ABSTRACT

The purpose of this correlational study is to understand the relationship between end-of-course grades as assigned by teachers and standardized end-of-course scores earned by students in Algebra, Geometry, Biology, Physical Science, and U.S. History courses at one virtual charter school in the State of Georgia. Pearson Product-Moment Correlation analyses were performed to determine if there is a statistically significant relationship between the numerical score earned in a course and the score earned on the required End of Course Test (EOCT) for subject of study. Separate Pearson Product-Moment Correlation analyses were run for students in disability subgroups. The Pearson Product-Moment Correlation analyses provide the researcher with a correlation coefficient, which allowed the researcher to determine the strength of the relationship. McNemar chi squared test on paired proportions were conducted to determine the association between pass rates (earning a 70% or higher) in the course and pass rates (earning a 70% or higher) on the required EOCTs. The results indicated that there is a significant positive correlation between passing a course and passing the mandatory End of Course Test. There was not a significant correlation for students with disabilities in the subjects of Ninth Grade Literature, Geometry and Physical Science. The results also indicated that there was a significant association between passing a course and passing the EOCTS for all subjects except Ninth-Grade Literature.

ACKNOWLEDGEMENTS

The author wishes to thank several people. I would like to thank my husband, Tommy Philipp, for his love, support and patience during the past four or so years it has taken me to graduate. I would like to thank my parents, grandparents and parents-in-law for their unending love and support. I would like to thank my children, Graysen and Charlie, for being my cheer team. I would also like to thank Natalie Hise, for her time, statistical knowledge and SPSS knowledge. I would like to thank all of my committee members, especially my chair Dr. V. O. Jones for her patience with me, as well as Dr. M. Griffin, for her unending support and guidance through both my Ed.S and Ed.D programs. I would like to thank the academic administrators at the school I am privileged to work at, for allowing me to use the schools data. Most of all, I would like to thank God for giving me the strength and endurance to accomplish my goal.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

Virtual school growth is a current trend in the United States. Parents choose virtual schools over traditional brick-and-mortar options for a variety of reasons, including safety, health, and religion. Most of the students who attend these schools are considered at-risk. The virtual school examined in this study qualifies for Title 1 funds that are dispersed by the federal government to schools where at least 15% of the school-wide population or 6,500 of the school's students are economically disadvantaged (U.S. Department of Education, 2009). Many of these students have been unsuccessful in traditional school environments, and virtual high school is the last alternative before dropping out of high school. Many students in virtual schools hold full-time jobs or are responsible for school-age siblings while their parents are at work. Despite the circumstances that brought these students to the virtual environment, they are still mandated by the states to take high-stakes graduation exams, in Georgia called End of Course tests.

Background

Mandated high-stakes end of course exams increase high school drop-out rates by as much as 1 percentage point overall in states with minimum competency exams and 2 percentage points in states where higher competency exams are required (Warren & Grodsky, 2009). Minimum competency exams require students to pass reading, writing, and math tests to graduate from high schools and are sometimes called Basic Skills tests. Higher competency exams require that students pass tests in a variety of subject in order to graduate from high school (Dee, 2003). Warren and Grodsky (2009) find that students who graduated from schools in states with high-stakes competency exams do no better in the labor market that those students who graduated from schools in states without high-stakes exams. Syverson (2009) contends that high-stakes EOCTs are geared toward middle-class students, and therefore, minorities, learning disabled, non-native speakers, and/ or impoverished students are left at a disadvantage. Syverson

(2009) also states that the reason high-stakes EOCTs continue to be utilized as a mandatory basis for graduation is that these tests generate more than $1.7 billion annually for private corporations that create and disperse the testing materials. For that reason corporations lobby legislatures to ensure their use across the country.

The current system is detrimental where it perpetuates labels for both students and schools. These labels are detrimental to student growth and to school identities. Students learn early if they are "good" at math or English or whether they are a "slow learner." Labeling of students is counterproductive to student instruction. Schools are also categorized as achieving or failing based on the outcome of end of course tests. Poor scores give schools a "failing" label that inhibits the task of improving a school's standing (Syverson, 2009). When schools fail—and at the expense of students' self- esteem--educational testing companies accrue additional profits selling tutorial and remediation products to states and school systems (Anyon & Greene, 2007).

Research shows that drop-out rates in states that require high-stakes tests is higher than the dropout rates in states that do not require such tests (Walden & Kritsonis, 2008). Students at risk of dropping out of high school are associated with five major demographics: poverty, race/ethnicity, family configuration, parental education, and limited English proficiency (Nowicki, Sisney, Stricker, & Tyler, 2004). The majority of high school drop-outs are already poor or belong to a minority group, but testing increases the likelihood that students in these demographic cohorts will not complete high school (Walden & Kritosonis, 2008). In fact, using these indicators, educators can predict with 80% accuracy which students will drop-out of school (Nowicki *et al*., 2004). McNeil, Coppalo, Radigan, and Heilig (2008) reveal that drop-out rates are highest for African American and Latino students. Thus, while other variables contribute to

the drop-out rate, studies suggest that standardized, high-stakes testing actually accelerates that rate.

One study, by Griffin and Heidorn (1996), suggests that failing minimum competency tests do not adversely affect minority students and the rate at which they drop out of high school in the State of Florida. But several studies counter that argument. One, a study by Madaus and Clarke (2001), indicates a direct correlation between high dropout rates and low socio-economic backgrounds in states where high-stakes tests are required for graduation. In this study, evidence from the National Educational Longitudinal Survey was used to determine if students who took high-stakes tests in eighth grade were more likely to drop out of high school by tenth grade. The evidence shows that those states that require high-stakes testing in eighth grade have dropout rates by tenth grade that are 4 o 6 percentage points higher than those that do not.

The study also discusses ethnicity and dropout rates. Madaus and Clarke (2001) look specifically at data from Texas after the state implemented the Texas Assessment of Academic Skills (TAAS). In 1991, Texas made passing the TAAS a requirement for successful completion of high school. In 1993, 40,000 sophomores dropped out of high school. The dropout rate for those 40,000 students was 25% Black, 23% Hispanic, and 13% White (Clarke, Haney, & Madaus, 2000). In 1989, before the implementation of TAAS, the number of Black graduates was close to the number of White graduates, with 76 percent of Whites and 74 percent of Blacks graduating. After implementation of TAAS, 70% of White ninth graders graduated in three years while only 50% of Hispanic and 50% of Black ninth graders graduated in three years (Madaus & Clark, 2001).

Problem Statement

The two-tiered foundation of standardized accountability is composed of two federal laws. No Child Left Behind Act, Public Law 107-100, of 2001 (NCLB), was introduced by the George W. Bush Administration in an attempt to hold students and schools accountable for meeting high standards of achievement. The American Recovery and Reinvestment Act, PL 111-155, of 2009 (ARRA), introduced by the Barrack Obama Administration, provided $25 billion to states in an attempt to underwrite adherence to NCLB; it gave an additional $4.35 billion to support the Race to the Top initiative (RTTI), a competitive program that provides select states with additional funding to improve the quality and consistency of the standardized tests they developed under NCLB. Georgia was one of the RTTI funding recipients (Barge, 2014).

Under both federal laws school districts must show that students are mastering the curriculum, which is summarily evolving to align with high-performance standardized tests. The federal guidelines thus regard students' performance on the tests as mastery of the curriculum. In many states, including Georgia, students must pass a battery of criterion-referenced tests. These states have put achievement tests in place to delineate whether students are promoted to the next grade or whether they graduate from high school. While teachers have always prioritized student performance, they are now under immense pressure from the federal government, state, administrators, and department chairs to raise standardized test scores. As a result, mandatory tests channel instruction to test preparation, resulting in a narrowing of the curriculum (Au, 2009). The new education initiatives of the Obama administration are trying to directly link teacher pay with student achievement scores, as part of the Race to the Top initiative (*American Recovery and Reinvestment Act*, 2009).

End of Course tests in Georgia are designed to gauge the knowledge that students gain in courses where state standards or objectives are being taught. The State of Georgia requires students to take End of Course tests in ten subjects: Algebra, Geometry, Math I, Math II, Biology, Physical Science, Ninth-Grade Literature, American Literature, U.S. History, and Economics. Students must take all of these tests and pass at least one in each core subject (math, English, science, and social studies) to graduate from high school. The tests in Georgia are designed so that students must answer half of the questions correctly in order to receive a passing score. These tests count for 20% of the course grade, a percentage that can affect students who might otherwise have passed a course and graduated high school. Though the state has established the reliability and validity of these tests, many students pass the course but fail the End of Course Test.

To reduce that outcome, the State of Georgia has developed a College and Career Ready Performance Index (CCRPI) as part of its use of RTTI funds. The CCRPI measures each school on a variety of factors, including student content mastery as determined by student achievement on standardized tests, in this case End of Course tests. As a result, many schools in Georgia find themselves with far more students getting credit for a course but not passing the End of Course Test associated with that course. This outcome has prompted researchers to ask if end of course grades accurately reflect the knowledge students have gained from the course or if they are inflated. This problem has been exacerbated in the virtual environment. This is most likely due to the fact that teachers do not have face-to-face contact with their students and the population virtual schools serve.

Purpose Statement

The purpose of this correlational study is to discover the strength of the relationship between passing a course and passing the EOCT associated with it in a virtual school environment. The variables are final grades, including exams, and standardized EOCT scores. These findings will allow virtual educators to better gauge whether the grading practices in virtual schools need to be addressed and changed to better reflect achievement as measured by the End of Course Test.

Significance of the Study

As educators, we can better serve students by targeting those most likely to drop out and providing them with counseling and access to test preparation programs. Since scores on high-stakes EOCTs such as those in Georgia are now linked to graduation, research is essential to understanding problems and adapting practices to them. Current research results suggest that school districts and states should target at-risk students regardless of where they go to school instead of condemning an entire school by removing funding. Monies should be awarded to schools that have large numbers of students who may be in danger of dropping out so that they can develop programs to keep students in school instead of punishing those schools where target groups fail to succeed on high-stakes testing. Because virtual school has such a high percentage of at-risk students and such low passing rates on standardized tests, it constitutes such an area where focus should be. This study will attempt to determine the strength of the relationship between students' end of course grades and EOCT scores at a statewide Georgia virtual school.

One of the stated objectives of both NCLB and RTTI has been to better prepare students for the labor force. However, Warren and Grodsky (2009) indicate that high-stakes exams do not benefit those who pass them in that they do not prepare students for the labor market better than

traditional curriculum, instruction, and testing. Research shows that high-stakes tests encourage teachers to change their pedagogy, with some using methods of lecture and rote memorization that have previously been shown to be ineffective compared to student-centered instruction. These methods leave students deficient in the critical thinking skills employers often seek (Au, 2009). NCLB also assumes that poverty rates are linked to low achievement on high-stakes tests instead of a lack of decently paying jobs (Anyon & Greene, 2007). Finally, test scores are subject to statistical manipulation, with states finagling scores in order to produce the desired pass rate (Yaffee, 2009), thereby obscuring who is successfully learning what. Martone and Sireci (2010) note that if gaps in curriculum alignment do exist, then measures should be taken to correct the problem. With tools such as the Effective Learning Program, educators can target students at risk of dropping out and provide support for them (Norwicki *et al*., 2004).

Research Questions and Hypothesis

Research Question 1: Is there a relationship between end of course grades and EOCT scores in Algebra, Geometry, Biology, Physical Science, Ninth Grade Literature, and U.S. History courses?

$H_0$1 There is no statistically significant correlation between student end of course grades (shown as percentages) and student EOCT scores in Ninth Grade Literature.

$H_0$2 There is no statistically significant correlation between student end of course grades (shown as percentages) and student EOCT scores in Algebra.

$H_0$3 There is no statistically significant correlation between student end of course grades (shown as percentages) and student EOCT scores in Geometry.

$H_0$4 There is no statistically significant correlation between student end of course grades (shown as percentages) and student EOCT scores in Biology.

$H_0$5 There is no statistically significant correlation between student end of course grades (shown as percentages) and student EOCT scores in Physical Science.

$H_0$6 There is no statistically significant correlation between student end of course grades (shown as percentages) and student EOCT scores in U.S. History.

Research Question 2: Is there a relationship between end of course grades and EOCT scores for students with disabilities?

$H_0$7 There is no statistically significant correlation between student end of course grades (shown as percentages) and student End of Course scores for regular education students and students with disabilities.

Research Question 3: Is there an association between pass rates (earning a 70% or higher) in the course and pass rates (earning a 70% or higher) on the End of Course tests required for that course?

$H_0$8 There is no association in passing (earning a 70% or higher) in the course and passing (earning a 70% or higher) on the End of Course Test required for that course.

Identification of the Variables

End of course grades

EOCT scores

Definitions

Academic achievement: For the purposes of this study academic achievement will be determined by the passing of the Georgia High School End of Course tests in Algebra, Geometry, Ninth Grade Literature, Biology, Physical Science and U.S. History. Students must pass one test in each of the academic areas (math, language arts, science, and social studies) in order to graduate from high school in Georgia.

Brick and mortar schools (BAM): For the purposes of this study a brick and mortar school refers to the traditional k-12 educational environment.

End of Course Test (EOCT): For the purposes of this study EOCTs will refer to the State of Georgia's required course examinations in American Literature, Ninth Grade Literature, Math I, Math II, Algebra, Geometry, U.S. History, Economics, Biology, and Physical Science. These tests count for 15% of a student's grade who entered high school prior to the 2011-2012 school year and 20% for students who entered high school in or after 2011-2012. (GA State Board Rule, 2006, 160-4-2-.13).

End of course grades or course grades: For the purposes of this study an end of course grade or course grade will be the teacher-awarded numerical percentage that appears after the class has ended. It will include the cumulative scores for all course work within the learning management system. End of course grades should reflect student mastery of the curriculum.

Learning management system (LMS): For the purposes of this study the learning management system (LMS) is the location of lessons, quizzes, and tests in the virtual environment.

Socio-economic status/economically disadvantaged: For the purposes of this study a student's \socio-economic status will be determined by the lunch status of the student. Students are considered to be from a low socio-economic background or economically disadvantaged if they qualify for free or reduced lunch.

Students with disabilities: For the purposes of this study a student with disabilities is a student who qualifies for special education services. There are two categories of students with disabilities: those who are taught in an inclusion setting and those who are taught in a contained classroom.

Virtual school: For the purposes of this study a virtual school is a public school in which students attend school from their homes, public libraries, or community centers using computers rather than in a traditional brick and mortar setting.

CHAPTER TWO: REVIEW OF LITERATURE

The purpose of this chapter is to review the literature regarding accountability, grading, high-stakes standardized testing, and virtual schools. This chapter begins with a theoretical framework, examining various theories as well as a biblical worldview. The next section discusses the history of accountability in the United States as well as the history of accountability in the State of Georgia. The following section reviews teacher grading policies and grade inflation in both secondary and post-secondary school. Literature on the impact of high-stakes standardized testing on students, schools, and teachers is examined in the fourth section. The final section of this review looks at current literature on virtual schools, with a primary focus on virtual k-12 schools.

Theoretical Framework

Student assessments in the constructivist model allow for exegesis, giving students the opportunity to show what they know (open-ended written assessments) rather than what they don't know (multiple choice assessments). The majority of instruction for end of course standardized tests is based on rote memorization for multiple-choice questions with no active learning and limited critical thinking involved. Constructivists theorize, however, that students retain more information when learning is active (Slavin, 1997) and that discovery learning increases student motivation, self-efficacy, and independence. Jerome Bruner, a leading proponent of discovery learning, states that such instruction allows the learner to acquire knowledge through experience and exploration. Bruner states that instructors merely provide the students with experiences for learning in a manner that allow a student to build on their prior knowledge (as quoted in Slavin, 1997). High-stakes tests do not promote the constructivist method of instruction because the tests are primarily multiple choice questions, and the attendant

curriculum must be detail-oriented so that students receive all of the exacting information needed to perform well on these assessments.

Those who defend the use of high-stakes tests often cite behaviorist theory, arguing that the reward of passing the standardized test motivates students. These theorists use Maslow's concept of "self-actualization" to explain that students rise to the occasion to perform well on high-stakes tests (Slavin, 1997). Critics argue that what behaviorist theories do not take into account is that students might have needs greater than those of educational performance, such as satisfying hunger or avoiding physical or psychological danger. These students do not have the energy to devote to learning and thereby do not achieve "self actualization."

Ryan and Weinstein (2009) also counter that high-stakes tests are not as motivating as behaviorists claim, and they use the self-determination theory as proof. The self-determination theory argues that using controlling measures (such as high-stakes tests) to change behavior only serves to de-motivate students to do well because the tests are often too difficult, with low scores representing negative feedback that prompt students to give up rather than repeat the action that caused the negative feedback, hence provoking students to drop out of school.

Biblical worldview implications

End of Course high-stakes testing was born out of the process/ mastery school of thought that sets as its principal idea a determination of whether or not students have mastered curriculum. Van Brummelen (2002) suggests that that approach may limit the ability to create critical thinkers who use reason and logic and who embrace the "spiritual, ethical, and aesthetic dimensions of learning." Van Brummelen (2002) contends that the world of standardized testing creates students who are "technically competent" but who "lack the commitments needed to foster a just and compassionate society." In advocating the integration of a Christian worldview

into the curriculum, he emphasizes the importance of creating learners who can think critically and who are aware of their impact on the world around them.

Accountability

Researchers argue about the origins of the accountability movement. Some researchers link the movement back to World War I, with the birth of the U.S. Army Alpha assessment test (Wineburg, 2004). Others link the creation of these tests to the launch of Sputnik by the Union of Soviet Socialist Republics (USSR), in 1957 (Garber, 2003). Whatever the origins, most agree educational accountability came into sharp focus with the passing of the Elementary and Secondary Education Act (ESEA), Public Law 89-750, 1965, and Public Law 90-247, of 1967. These two key pieces of legislation within Lyndon B. Johnson's War on Poverty were established to improve educational opportunities for the poor by disbursing Title 1 funds to state educational agencies, which then distribute the monies to local agencies. Title 1 funds are to be designated to schools with disadvantaged students or schools with high poverty rates.

From 1965 to 1985, the United States saw a doubling in the amount of federal funds given to elementary and secondary schools as well as an increase in spending on education. During this same period elementary and secondary schools saw an increase in federal involvement that focused on equity and access rather than on student achievement and school accountability. This changed with the 1983 publication of *A Nation at Risk: The Imperative for Educational Reform* (McGuinn, 2006).

A Nation at Risk and Improving America's School Act

In 1983 the National Commission on Excellence in Education issued *A Nation at Risk*, a study of schools across the country that brought to light the ineffectiveness of ESEA's Title 1 funds. The commission found that the monies were not leveling the playing field for

impoverished students and that American children lagged behind other countries in achievement, a theory that linked the nation's growing economic issues to students' lack of preparedness for the workplace. Its authors advocated reform, calling for the federal government to step in to provide guidance on national education and to provide funds to states and local school districts to implement federal guidelines. Some suggestions from the commission included increased teacher pay and strengthened curricula and standards. The report also argued that achievement tests be put into place to make sure students were being taught basic skills. Within two years of the report, at least some states began implementing the reforms suggested without any federal incentives (McGuinn, 2006; Sunderman, 2009).

By 1993, however, the William Clinton Administration began offering federal money as an incentive for all states to adopt high content and high performance standards. Improving America's Schools Act, Public Law 103-392, of 1994, reauthorized ESEA and provided more federal money to states that implemented high standards and the measures to assess effectiveness, i.e., high-stakes tests. Despite the offer of money, not many states had fully implemented measures for assessing the implementation of content and performance standards by 2001 (Sunderman, 2009).

No Child Left Behind and Race to the Top

The George W. Bush administration took the work of the Clinton Administration to the next level with NCLB, which reauthorized ESEA, and aimed to increase accountability of the nation's schools to provide a world-class education to all students. As a result, it required that states be accountable for the results of tests by reporting the data from them (Byrnes, 2009).

This mandate has led to increased scrutiny and criticism of NCLB and the high-stakes tests associated with it. Central to that criticism has been the fact that these scores affect whether

a school meets adequate yearly progress (AYP). As defined by the U.S. Department of Education, schools, districts, and states are issued report cards based student achievement, attendance, and graduation rates, which are reported on state websites for the public to review. If a school fails to make AYP for three consecutive years, the school may be subject to a re-organization by the state (Katsiyannis, Zhang, Ryan, & Jones, 2007). In turn, the increased pressure on schools results in increased scrutiny of high-stakes tests.

President Obama's Race to the Top program also reauthorized ESEA and NCLB. The Obama initiative aims to reform the lowest performing 5% of schools in the nation and offered large monetary grants to states that applied and met very specific requirements. Schools' use of pay-for-performance is one of these requirements. Another is teacher evaluations, which tie student achievement on high-stakes end of course tests to teacher pay. This method of evaluation is supposed to act as leverage to increase student achievement (U. S. Department of Education, 2009; Maxcy, 2011).

In March 2012, Georgia successfully applied for a waiver from some of the mandates of NCLB, one of 10 states to receive such a waiver. The provisions of the waiver require that Georgia must identify and concentrate on reforming those schools identified as Title 1 priority and focus schools. These schools are identified based on achievement data and graduation rates. Georgia developed the College and Career Ready Performance Index (CCRPI) for state accountability purposes (Barge, 2013).

Accountability in Georgia

At one time, Georgia required only a minimum of Carnegie units and no high-stakes exit exam to graduate from high school. In 1983, however, Georgia implemented the Basic Skills Tests (BST) in response to the *A Nation at Risk* requirements. The Basic Skills Tests consisted

of an assessment of mathematics, language arts, and writing, and students had to attain a minimum passing score to graduate from high school. The requirement to pass the BST remained in effect until 1994, when it was replaced by the Georgia High School Graduation Test (Georgia Department of Education, 2013).

Those students who entered ninth grade after the 1990-1991 academic year were required to take the Georgia High School Graduation Test, which measured competency in English/ language arts, mathematics, and writing. Students who did not pass these tests were unable to graduate from high school with a general education diploma. Georgia added an assessment in social studies in 1993 and an assessment in science in 1994, and all five tests had to be passed in order to graduate high school. These tests assessed student knowledge of the Quality Core Curriculum learning objectives that had been in place at the time.

Starting in 2005, Georgia introduced a new set of learning objectives called the Georgia Performance Standards. Georgia Performance Standards in English, writing, and science were introduced first, followed by social studies, in 2007, and mathematics, in 2008. Students first take the Georgia High School Writing Test in October and the Georgia High School Graduation tests in March of their junior year. They are allowed six tries to attain a passing score on the Georgia High School Writing Test and five tries to attain a passing score on the individual Georgia High School Graduation tests during their junior and senior years (Georgia Department of Education, 2013).

In an attempt to meet the mandates of NCLB, Georgia created End Of Course tests, commonly referred to as EOCTs, in Ninth Grade Literature, American Literature, Algebra, Geometry, Biology, Physical Science, U.S. History, and Economics/Business/Free Enterprise.

These tests were created not only to ensure that courses in Georgia were rigorous, but also to provide data as a means of studying effective instruction and student achievement (Cox, 2006).

Between 2001 and 2009 students entering the ninth grade in Georgia were required to take both the Georgia High School Graduation tests in English/Language Arts, Mathematics, Science, and Social Studies and the End of Course tests in the eight core subjects listed above. Those students who entered high school in 2010 could exempt the Georgia High School Graduation tests provided they had taken and passed an End of Course Test in the same subject. For example, if a student had taken and passed either the End of Course Test in Ninth-Grade Literature or American Literature, the student was not required to take the Georgia High School Graduation Test in English/Language Arts. Students entering ninth grade after 2010 are required to pass one End of Course Test in each of the core academic areas of English/Language Arts, Mathematics, Science, and Social Studies in order to graduate from high school. EOCT scores were calculated as 15% of students' course grade prior to 2010 and 20% for those entering high school in 2011 and after (Georgia Department of Education, 2013). Teacher-issued grades made up 85% of the course grade prior to 2010 and 80% thereafter.

Grading

A hodgepodge of different measures, including aptitude, effort, punctuality, and behavior, determines what mark a student receives in a course (Guskey, 2011). A bright student who scores well on course and state tests but who consistently turns in work late may receive a mark similar to an average student who turns in everything on time but who scores poorly on those tests (Guskey, 2006). Grading systems all over the country are so subjective that is becomes difficult to gauge what a student has actually learned and how good a student has

become at "playing the game," particularly when they can gauge teachers' varying grading methods or policies and adapt their behavior to them.

A study by Randall and Engelhard (2010) of 516 school teachers reveals a strong correlation when comparing numerical and alphabetical grades. The researchers point out that teacher grades are not completely based on student achievement. Many teachers consider such measures as ability, behavior, and effort to determine student grades.

Post-secondary and Graduate Grade Inflation

Grade inflation, whether at the college or high school level, is an "increase in grades without a concomitant increase in achievement" (Ziomek & Svec, 1997). The seemingly endless variables that contribute to the awarding of course grades is reflected in a body of literature that focuses on a variety of post-secondary settings. That literature suggests important guideposts for high schools, particularly in terms of documenting grade inflation over time and the impending link between student achievement on high-stakes tests and teacher pay.

At the undergraduate college level, the number of As being awarded has dramatically increased since the 1960s (Kamber & Biggs, 2003). The pre-1960 meaning of grades--A for outstanding, B for above average, C for average--simply no longer apply. Students have come to think of a B not as an above average grade, but as average grade that the majority of them can expect.

Examining the evolution of grades in 200 four-year colleges between 1940 and 2009, Rojstaczer and Healy (2012) find that the awarding of grades stayed consistent from 1940 to 1960, but that the number of As awarded after 1960 increased. The percentage of As awarded since 1960 has equaled 43%, an increase of 28% since 1960. The evidence shows that grades were awarded on a true bell curve prior to the 1960s, when Cs were most often earned. The

researchers' evidence shows a distorted bell curve by the 1980s, when Bs were most often earned. The bell curve completely disappears by the early 2000s, when As became the frequently awarded grade, making the "curve" no longer a curve but a distinct ski slope. The two researchers contend that the data suggest an increasing lack of oversight by the universities' leadership. Rojtaczer and Healy (2012) also note that grade inflation at the post-secondary level has impacted professional and graduate schools as well and thus employers' ability to use grades to identify qualified applicants. Indeed, they say that grade inflation at the post-secondary level has become so rampant that grades no longer reflect a student's actual performance in a course (Rojstaczer & Healy, 2012).

The reason post-secondary instructors and professors have inflated grades may be revealed in two studies. In the first, by Kamber and Biggs (2003), the researchers contend that instructors and professors inflate grades not as a product of learning but to keep students happy so that the students will evaluate them favorably, thus keeping administrators at bay. Professors and instructors also may be motivated by money, as their evaluations are frequently used by universities to determine pay increases, promotion, or whether a visiting professor is asked back to teach the course (Germain & Scandura, 2005). This finding weighs heavily on the future, as RTTI requires high school achievement to be tied to teacher pay.

Kezim, Pariseau, and Quinn (2005) conducted a similar longitudinal study, analyzing the grade point averages of business school students over a 20-year period and looking specifically at the grades awarded to students in courses where tenured and adjunct professors taught the same course. The researchers find that the course grades and overall grade-point averages of students taught by adjunct or non-tenured professors were higher than those taught by their tenured peers (Kezim *et al*., 2005), perhaps because the jobs of non-tenured and adjunct

professors are more tenuous and open to dismissal. To ensure job security, therefore, these men and women may have similarly tried to keep administrators at bay by inflating grades.

Grade Inflation in High School

To adequately remark on grade inflation over time not only requires longitudinal study but also a basis of comparison. To establish that comparison, a number of researchers have compared course grades to scores on the American College Testing (ACT), Scholastic Achievement Test (SAT), and Advanced Placement (AP) exams. What researchers have found is that an A is no longer an A in a course but merely grade inflation.

A study by Woodruff and Ziomek (2004) compares grade point averages with ACT scores among high school students between 1991 and 2003. The researchers use self-reported data from students in 23 classes to determine students' grade point averages and compare those averages to the students' most recent ACT scores on record. Woodruff and Ziomek reveal that high school grade point averages inflated approximately 6% over a 13-year period. The researchers maintain that grade inflation is especially prevalent for students with low ACT scores, with limited amounts of inflation for students with higher scores (Woodruff & Ziomek, 2004). In a more recent study, Goodwin (2011) finds that the average grade point average has risen from 2.80 to 3.04 while in the same period the average ACT score rose only slightly, from 20.04 to 20.55 on the math portion and from 20.22 to 20.44 on the English portion. Still, critics call into question the accuracy of the grade point average data used in these types of studies because students self-report their own grade point averages.

Using the Scholastic Aptitude Test (SAT) instead of the ACT, Godfrey (2011) compares exam scores to grade point averages and had findings similar to those of Woodruff and Ziomek (2004) and Goodwin (2011) for the ACT. Godfrey calculates that the average grade point

average of 2.64 in 1996 had risen to 2.90 by 2006, but SAT verbal and mathematics scores did not show a comparable increase during the same time period. In fact, the average SAT score on the verbal section of the test dropped from 497 to 495 in 2006.

In that same study, Godfrey (2011) also compares Advanced Placement (AP) course scores with AP test scores in five subjects at five schools. The researcher finds that there is no consistency in teacher or school grading policies as they relate to AP test scores (Godfrey, 2011).

Potential Reasons for Grade Inflation in High Schools

As grade inflation became apparent after these studies were published, researchers have tried to determine the reasons for grade inflation, and what they have discovered points to federal laws that require schools to raise scores on high-stakes standardized tests, increase the pass rate, and by extension, decrease the dropout rate. As schools across the country feverishly try to make AYP or jump over other hurdles put into place as a result of these laws, policies and practices are being instituted across the country that unfortunately result in grade inflation.

Cox (2011) closely examined one school district and finds that the policies it instituted as a result of the legislation has led to significant patterns of rising grades. High schools in this district use strategies such as course "alike" grading policies, a minimum of 50% to represent an F, a test retake policy, and a late work acceptance policy. Course "alike" policies are created by teachers of the same course to determine "what counts" for credit in a course, i.e., "alike" assignments and grading methods (Cox, 2011). The district reports that it had instituted these grading practices in an effort to "keep hope alive" by reduce the number of course failures and increasing the numbers of graduates. The policies do indeed reduce failures, but the quality of learning may have been lost in the process.

Other policies used by schools that contribute to grade inflation include "minimum grading policies" or "zeros aren't permitted" (ZAP) programs. Schools using minimum grading policies often set a floor on the score students can receive in the first quarter to make earning an overall passing score for the semester possible. For example, mandating a minimum of a 55 or 60 for first semester makes passing for the year attainable while the actual 30 or 40 the student earned does not. Other schools go so far as to not allow the entering of zeros in the grade book for missing work or requiring students to attend a homework café during lunch time or after school to make up the assignments. These programs are typically put in place by schools with large at-risk populations in an effort to boost self-efficacy and self-motivation, but critics argue that these programs falsely reward students and result in grade inflation (Carifio & Carey, 2010). Moreover, research suggests that low grades do not encourage students to work harder but instead prompt students to withdraw from learning (Guskey, 2011).

As the number of high school graduates that pursue post-secondary education increases--51% in 1975 to 69% in 2007--college readiness has become an increasing concern for both educators and policy makers (Porter & Polikoff, 2011). Today, a staggering number of high school graduates who appeared to have the requisite grades to be accepted into college have entered college needing remediation in reading and mathematics as a result of inflated grade point averages. Howell (2011) reports that of the nearly 400,000 students the California State University system admits each year, between 40% and 50% of these students need remediation in reading and/or math. A report by the Manhattan Institute for Policy Research indicates that only about 32% of the high school graduates it studied are prepared for college. The numbers of Black and Hispanic students who leave high school prepared for college is significantly lower (Greene & Forster, 2003).

Grades and fairness

One would expect higher grades, whether by inflation or in actuality, to result in students' perception that their teachers have been fair, but studies show that this is not the case. Other variables, particularly the time and attention of the teacher, influence student perception. Together with the studies that show a lack of readiness for college, these re[prts suggest that perhaps district, school, and teacher policies should restrict the avenues that lead to grade inflation.

Gordon and Fay (2010) conducted an exploratory study to determine what comprised students' perception of fair grading. The researchers compare two groups: 193 college students in three management courses received course credit for completing the questionnaire, and 473 organic chemistry students received no credit for completing the questionnaire. The results show that student perception of fairness is related more to the instructors' efforts to assist students to do well in a course, i.e., offer study guides, practice tests, or review sessions, rather than grade curving or dropping of low grades (Gordon & Fay, 2010).

Universities understand the implication of grade inflation both at the high school and undergraduate levels and so have changed the ways they admit students. Wongsurawat (2009) evaluated grade inflation and its impact on U.S. law school admissions. His findings are so revealing that law schools now rely more on standardized tests than on grades to determine admission eligibility of law school candidates (Wongsurawat, 2009). This has increasingly been the policy of schools throughout the country, weighing standardized tests such as the SAT, ACT, or Graduate Record Exam (GRE) more than student grade point averages. The irony is that elementary and secondary school standardized tests are what led to the grade inflation to begin with.

Standardized Tests

Standardized tests are considered a necessary evil in this time of ever increasing accountability. In most cases the standardized tests have become high-stakes for students, teachers, and schools, elementary and secondary. This section discusses the difficulty levels of standardized tests used in schools to measure achievement, the negative perceptions of these tests, and the relationship between the existence of these tests and the high school dropout rate. The impact of these tests on curriculum and instruction and on teacher and student motivation are also reviewed.

Difficultly level of standardized tests

Researchers working with the Achieve Incorporation do not believe that the difficulty level of graduation tests is an unreasonable expectation of high school students (Achieve, Inc., 2004). Researchers looked at graduation exams from six states and examined the complexity of the math and language arts materials presented. The researchers argue that the material tested is perfectly reasonable and that students should not have problems successfully passing the tests. In fact, the language arts material tested is what the ACT expects of students in the eighth or ninth grades. In math, students are only required to know what other countries require of students in the eighth grade, based on the International Grade Placement scale developed by Michigan State University. As a result, the researchers suggest that high school tests should be strengthened over time to increase student learning, not only because 53% of students entering college require remedial courses in English and math, but also because more than 40% of those students who do require remedial courses will not earn a 4-year degree (Achieve, Inc., 2004).

High-stakes Tests' Effect on Student Performance

In theory, high-stakes tests are designed to measure student learning and increase learning over time by making each sequential test harder than the last, much as the Pre-SAT presages the SAT. Research has shown conflicting results on the effect of high-stakes tests, however.

On the one hand, Amrein and Berliner (2003) argue that high-stakes tests are negatively affecting learning. Focusing on 18 states that use high-stakes course and graduation tests, they compared student scores on ACT, SAT, AP, and National Assessment of Educational Progress (NAEP) exams. The research shows that in 18 states, learning stayed the same or decreased in all but one case after high-stakes tests were implemented. Another study indicates that only in states where the tests were poorly designed or the testing materials had been compromised is there a low correlation between learning and test scores (Greene, Winters, & Foster, 2003).

Berliner (2009) also reports on another study where SAT scores are compared in states with and without high-stakes exit exam requirements. The study finds that students in states without high-stakes exit exams out-perform those states where high-stakes exit exams are required. The study reasons that this is due to the narrowing of curriculum in states that require high-stakes exit exams. The narrowing of the curriculum stifles students' ability to reason, which is one of the skills tested on the SAT.

On the other hand, there are a number of studies that refute the claim that standardized tests decrease learning, at least among those at the top of their classes. Re-examining the Amrein and Berliner (2002) study and using the same NAEP data, Braun (2004) compares the 18 states with high-stakes tests to those states without high-stakes tests. He reports that students in high-stakes states scored better on the NAEP math test in the fourth and eighth grades. States with

high-stakes tests performed better than states with low stakes tests on the NAEP across the nation. Greene *et al*. (2003) also notes a high correlation between states' scores on college entrance and credit tests and high-stakes course and graduation tests. These researchers find gains on the Florida Criteria Assessment Test and scores on the Stanford 9 test were positively correlated. Therefore, the researchers conclude, high-stakes tests can be accurately used as a tool to measure student performance on college entrance and credit tests (Greene *et al*., 2003). Even Berliner, in a separate study (2009), shows that high scores on end of course tests at least marginally increases the probability of a completed post-secondary program (Berliner, 2009).

Relationship between high-stakes tests and graduation/ dropout rates

A strong statistical relationship exists between high-stakes as a graduation requirement and the high school dropout rate. According to Clark *et al*. (2000), the 10 states with the highest dropout rates (or lowest rates of high school completion) are states that require students to successfully pass a high-stakes graduation test. The 10 states with the highest rate of high school completion (or the lowest dropout rates) do not require a passing score on the state exam in order to graduate from high school.

This Clark *et al*. study (2000) illuminates that it is not so much the administration of a test but rather the use of the test as a gauntlet for graduation that affects performance. Presumably, one could call that "pressure," as one research study does. Nichols, Glass, and Berliner (2006) created a pressure index to measure the impact of high-stakes testing. The index was used to find the correlation between pressure and test results in 25 states. The researchers find that pressure to pass high-stakes exams causes a negative effect on students and that students are more likely to be retained or drop out in states that had a high-pressure score.

Jonsson (2001) reports that the number of students who did not complete high school in North Carolina rose 17% from 2000 to 2001 as a result of the implementation of high-stakes, high-pressure tests. According to Amrein and Berliner (2003), 88% of those states with graduation tests have higher dropout rates than those states that do not have graduation tests. Amrein and Berliner (2003) also report that in those states with mandatory graduation exams, dropout rates are 4% to 6% higher than in those states without them. Similarly, a study by Jacob (2001) indicates that students in the lower 25th percentile are more likely to drop out of high school in states that require graduation exams than their counterparts in states where high school graduation exams are not required. Haney, Maduas, Abrams, Wheelock, Miao, and Gruia (2004) conducted a longitudinal study with similar results, studying ninth graders who did or did not graduate in four years between 1971 and 2000. They find that the graduation rate dropped from 77 percent in 1971 to 67 percent in 2000. The researchers attribute this decline to minimum competency tests, an increase in academic standards, and an increase in graduation requirements.

Not all researchers agree that graduation test implementation creates negative results. Greene and Winters (2004) show that implementing graduation exams has had no significant effect on graduation rates. Greene and Winters use two different methods of calculating graduation rates, and both methods yield the same results: Graduation rates do fluctuate from year to year but are not affected by the implementation of graduation exams. Using one of two methods of calculation, the researchers find that teacher-to-student ratios and spending-to-pupil ratios also have no significant effect on graduation rates. The other method shows a small increase in graduation rates when numbers of students to teachers increase. In the discussion, the authors mention that graduation exams do keep some students from graduating. Consequently, schools are striving to produce better students who are capable of passing exams, and they are

working to target those students who may be in danger of dropping out before they make it to graduation exams (Greene & Winters, 2004). Likely these are not the students who do benefit from high-stakes tests, as discussed in the previous section. Instead, they are likely to be non-college preparatory tracked students and/or at-risk.

High-stakes Tests and Student Motivation and Learning

Legislators believe that by creating higher standards, states, schools and students will rise to the challenge and work harder to achieve ever higher percentages of passing scores on high-stakes graduation exams. However, as high-stakes tests were implemented and graduation rates decreased, especially among target populations, education researchers began to look at the data to see if students were more motivated to succeed after the tests were implemented. Most studies show no direct link between increased motivation and high-stakes tests. Indeed, they show high-stakes testing adversely affects student motivation and learning (Paris & McEnvoy, 2000; Amrien & Berliner, 2003).

For example, Paris and McEnvoy (2000) find a correlation between high-stakes tests and student motivation. In lower grades students are more likely to do their best on tests, believing in the validity of the tests. In high school, however, students are not as accepting of the validity of high-stakes tests because they did not believe the tests reflected the curriculum taught. As a result, high school students are more likely to participate in counterproductive test-taking behaviors, such as cheating, loafing, and discounting the value of tests. Students who do not score well on high-stakes tests are more likely to participate in the poor test-taking strategies listed above and their motivation levels significantly decrease.

Amrien and Berliner (2003) collected SAT, ACT, AP, and NAEP reports in 18 states that require exit exams. The researchers see that in those states that require high school graduation

exit exams, student motivation and learning do not increase. In fact they find that in some states, the scores on these national tests actually decreased after the implementation of graduation tests, refuting studies such as those by Bruan (2004) and Greene *et al*. (2003). The Amrien and Berliner study (2003) indicates that in the years after implementation of high-stakes graduation exams, SAT scores in seven states increased while scores in the other 11 states decreased. Therefore, the Amrien and Berliner study indicates, "high-stakes testing policies have no systematic effects on learning" (p.35), neither increasing nor decreasing learning.

One recent study also shows that increasing student test scores has no correlation to student earnings after high school. The study suggests that students would have benefited more by increasing their nonacademic skills in leadership, communication, and responsibility than by taking high-stakes tests (Berliner, 2009). However, if the studies by Bruan (2004) and Greene *et al*. (2003) are true and high-stakes do help certain students do better on college entrance exams such as the SAT and ACT and they are among the percent who do make it through college, then their earnings will be higher than those of students who drop out of high school or who earn a high school diploma. Moreover, this supports this essay's contention that students should be targeted, helped, and supported so that they do not drop out of high school.

Teacher Motivation

Students are not the only group feeling undue pressure by schools. High-stakes testing also affects high school teachers' motivation and morale in classrooms as well. Pedulla, Abrams, Madaus, Russel, Ramos, and Miao (2003) conducted a survey of teachers from states with and without high-stakes tests. Survey questions focused on test preparation, content and mode of instruction, and teachers' perceived value of the tests. Teachers in high-stakes states were more likely to spend more time on test preparation than teachers in states without high-

stakes tests. The former spent more time on instruction of material tested than on non-tested material. The survey shows that technology was used less by teachers in high-stakes states than by teachers without high-stakes testing. Also, teachers in high-stakes states were less likely to pursue field trips or enrichment activities. The survey reports that teachers do not think the tests are worth the time and money involved, but they do believe that the tests bring to light important educational issues.

Abrams (2004) supports this work. Abrams studied the results of a 2001 National Board Of Educational Testing and Public Policy survey to which 4,200 elementary and secondary education teachers responded. He looked specifically at 167 teachers from Florida who responded, and he compared their responses to teachers in other states with high-stakes testing. In both cases he finds that teachers believe that quality teaching and learning has been replaced by test preparation. In fact, some teachers surveyed in the study refer to their schools as "test prep factories." Abrams (2004) finds that 62% of the Florida respondents greatly increased the time spent on instruction in tested areas, while 67% decreased the amount of time spent on instruction in non-tested areas. Eighty percent of surveyed teachers in Florida state that they feel pressure from their district superintendent to raise scores, and 48% state they feel pressure from their building principal (Abrams, 2004).

In-depth qualitative research has produced similar findings. Mastropieri *et al*. (2005) performed four case studies in a long-term qualitative investigation of co-teaching in science and social studies. One major challenge faced by these teachers was the presence of high-stakes testing. The authors reveal that co-teachers report increased pressure by high-stakes tests, a lack of differentiation in their classes, and a limit to and modification of the presentation of material in order to meet time constraints due to high-stakes testing.

A qualitative study by Finnigan and Gross (2007) looked at teacher motivation in relationship to accountability measures. Ten teachers from schools on probation in Chicago were interviewed about their experiences with high-stakes testing. The teachers report increased pressure and believe expectations are too high while others feel discouraged by the negativity of the seemingly endless discussions of projected failures. Stillman (2009) conducted three case studies of elementary teachers who were preparing students in the early years of NCLB in an "underperforming" school that has a large number of Spanish-speaking English language learners. The teachers in this study report that teaching the standards is not the issue; instead, they believe that the standards do not embrace the students' native language and culture. A study by Hunt *et al*. (2009) focuses on the Urban Research Team at Southern Illinois University's work on teacher satisfaction among schools that did not make AYP. Based on open-ended responses to interview questions, the authors report that teachers feel rushed and unable to provide time for enrichment activities and in-depth curriculum for students due to the demands of high-stakes tests.

Pressure for students to achieve on high-stakes end-of-course tests could be a factor in the high teacher attrition rate in poor urban schools. Borman and Dowling (2008) found that teachers who work in urban schools are more likely to leave the teaching profession than teachers who work in suburban schools. The teachers who leave the profession are usually more qualified than those who choose to stay. Schools with high percentages of poor and minority students are also more likely to have a higher teacher attrition rate than those schools with low percentages of poor and minority students. The narrowing of curriculum also has had a negative effect on teacher retention rates (Crocco & Costigan, 2007). This is mainly due to the pressure teachers feel concerning student achievement on the high-stakes tests. Schools narrow

curriculum in an effort to raise achievement, and teachers excitement for their profession wilts as a result of rote curriculum and instruction.

Indeed, the pressure is so great that teachers in low performing schools in Texas report using questionable practices while administering the Texas Assessment of Academic Skills. Some of these questionable practices include giving reviews during the test, giving hints, pointing out mismarked answers, and directly pointing out correct responses (Paris & Urdan, 2000).

Impact on Curriculum and Learning

High-stakes tests have a negative impact on curriculum and instruction both at the elementary and high school levels. High school social studies and science teachers are often at a greater disadvantage than their colleagues in math and language arts because of the decreased time spent on social studies and science curriculum at lower grades.

Au (2009) shows evidence of studies where the narrowing of curriculum occurs among social studies educators whose students must take high-stakes tests. The studies cited in Au (2009) show teachers change pedagogy practices from high-level inquiry based instruction to low-level rote memorization. Au points out that in many U.S. school districts, the majority of instructional focus in K-5 schools is on reading and math tests and that these schools are limiting, if not eliminating, social studies instruction.

A later case study by Wills and Sandholtz (2009) finds that teacher decisions are greatly constrained by the pressures of high-stakes tests even in subject areas that are not being tested. In their fifth-grade classrooms the teachers in this study allowed instructional time allotted for social studies to be used for math and language arts instruction because math and language arts

have high-stakes tests and social studies do not. This shows that the effect of high-stakes testing impacts more classes that just those that require high-stakes tests.

Pace (2008) interviewed three fifth-grade teachers in California who contend that social studies instruction is marginalized in order for schools to concentrate on math and language arts instruction. Similarly, a study by Nichols and Berliner (2008) included an elementary teacher in Colorado who reports that she no longer uses the science unit where she hatched baby eggs in her classroom because she was told to focus more time on reading. She reports that she also has cut the community outreach activities she used for citizenship curriculum.

The presence of a high-stakes test in one subject can cause teachers to shut out subjects and curriculum that are not tested. A study by Gerwin and Visone (2006) presents a narrative of two social studies teachers and discusses the impact of high-stakes testing on each. Both teachers kept a journal of their daily activities: one teacher did not have an End of Course Test, and the other did. The researchers explain that the teacher with the End of Course Test focused more on drill and test preparation while the second teacher engaged in more thought-provoking activities.

High school teachers have similar issues. Journell (2010) finds that six government high school teachers from three schools in Chicago did not allow significant time to cover the 2008 election in government class due to the time constraint of preparing students for the high-stakes end-of-course test in government required for graduation. At the high school level, the social studies curriculum has been reduced to memorization of a limited, prescribed curriculum. The social studies curriculum has become more of a trivial, fact-based memorization than a student-centered exploration of events (Au, 2009). Thus, schools are "teaching to the test" in order to attempt to save their "failing" school. Yet standardization of content has not shown great

advantages in improving test scores. Instead, the system of testing and re-testing promotes rote memorization instead of the critical thinking skills that students will need once they leave the school environment and enter the "real" world.

Wills (2007) presents similar findings in California. The study notes the lack of social studies instruction in several schools in southern California identified as low performing. This article is evidence of the absence of social studies curriculum at the lower grades, which directly impacts the lack of knowledge high school teachers' encounter when these low performing students reach high school.

Narrowing of the curriculum also occurs in different ways in low- and high-performing schools. Paris and Urdan (2000) find that teachers in low performing schools spend more time on test preparation than teachers do in high performing schools. The study found that some teachers begin reviewing for the high-stakes test two months in advance of the exam to the exclusion of content deemed non-essential--i.e., not on the test.

Curriculum Alignment

The presence of high-stakes testing has not only narrowed and eliminated course studies material, it also has dictated curriculum alignment. As part of federal requirements, states must prove that state assessment tools align with state standards (Martone & Sireci, 2010). However, it is difficult to prove that classroom instruction aligns with the state-mandated curriculum.

Martone and Sireci (2010) reviewed three methods of evaluating and aligning curriculum. Curriculum alignment not only refers to classroom alignment of instruction, activities, and assessment but also alignment of skills and knowledge from year to year and from class to class. Some alignment methods include Webb's Alignment Method, Achieve, and the Surveys of Enacted Curriculum Method. Webb's method looks at how closely aligned the assessments are

to the objectives. The Achieve method does the same but on a micro level, looking at each item to assure alignment to the objectives. The Surveys of Enacted Curriculum method helps educators see a connection between what is taught in the classroom and what is assessed on the test (Martone & Sireci, 2010).

High-stakes Tests Effects on Poverty

From the 1960s onward, ESEA and NCLB have been seen as anti-poverty measures in that they ostensibly attempted to eliminate poverty by increasing educational opportunity and achievement (Anyon & Greene, 2007). As a result of this vision of education, many states instituted high-stakes tests to determine their progress. Now, researchers are finding that obtaining a high school diploma does not guarantee a job for recipients nor does the testing and subsequent diploma decrease the poverty rate as advertised.

Moreover, wage differentials continue to exist regardless of attainment of a diploma: Anyon and Greene (2007) find that male high school drop-outs continue make more money than female high school graduates just as they did before high-stakes testing was put into place. The research also shows that the number of welfare recipients who have high school diplomas has increased from 42% in 1979 to 70% in 1999, giving lie to the idea that a high school diploma will lift recipients out of poverty.

Although job training was introduced to reduce poverty rates as early as the 1990s, federal legislation that mandated such training failed to include job creation. In the end, students prepare for high stakes tests in hope of securing jobs that do not exist (Anyon & Greene, 2007).

High-stakes tests and students with disabilities

Students with disabilities struggle to pass the Georgia High School Graduation tests (GAOSA, Report Card). This is not a phenomenon that affects only Georgia. The Virginia

Department of Education revealed in 2005 that 72% of students with disabilities compared to 86% of all students in that state passed the state Science Standards of Learning Test in 2004 (Mastropieri, Scruggs, Norland, Berkeley, McDuffie, Tornquist, & Conners, 2006).

However, according to one study by Ysseldyke, Nelson, Christenson, Johnson, Dennison, Triezenberg, Sharpe, and Hawes (2004), little empirical evidence is available on the connection between high-stakes testing and graduation rates for students with disabilities. Ysseldyke *et al*. (2004) gathered what empirical evidence there is and studied newspapers to assess the impact of high-stakes tests on students with disabilities. The researchers report that the evidence they found was neither overwhelmingly positive nor negative. High-stakes exams are used to determine grade promotion for all students, including those with disabilities. Ysseldyke *et al*. (2004) suggest more research should be conducted to determine if these tests negatively affect students with disabilities compared to their non-disabled peers.

Not all the effects of NCLB are so ambivalent. Christenson, Decker, Triezenberg, Ysseldyke, and Reschly (2007) surveyed educators from 19 states with high-stakes tests and find that most educators believe the new requirements have had a positive effect on teaching and learning for student with disabilities. Survey respondents report that students with disabilities are monitored more closely as a result of the implementation of high-stakes tests. Ysseldyke *et al*. (2004) finds that teachers are spending more time teaching the standards to be tested and less time teaching information that will not be tested. Christenson *et al*. (2007) reaffirms that teaching focuses more on standards. In the survey that they administered, respondents report sensing that leaving out information that will not be tested was the only negative effect of high-stakes tests.

As a result of NCLB and Individuals with Disabilities Act (IDEA) Public Law 94-142, the number of students with disabilities who participate in high-stakes tests has increased (Schulte, Villwock, Whichard, & Stallings, 2001; Katsiyannis *et al*., 2007). With increased participation has come an increase in the number of disabled students who achieve proficient scores (Schulte *et al*., 2001). While overall scores have increased, the percentage of students who are not successful on high-stakes tests is still large in some areas. Numerous studies have been conducted on strategies to improve the high-stakes test scores of students with disabilities.

Virtual Schools

The National Center for Education Statistics defines "virtual school" as any public or private school that offers only virtual courses and that does not have a brick-and-mortar facility to allow students to attend classes on site (NCES, 2013). "Virtual education" is defined as instruction in which the instructor and students are separated by time and space. This instruction can either be coursework presented in an online environment accessed at the student's leisure or as actual instruction accessed in real time by the student.

Virtual schools are divided into five different types of administrative structures: statewide supplemental programs, district-level programs, single-district cyber schools, multidistrict cyber schools, and cyber charter schools. There are approximately 227 virtual schools of different varieties in 47 states. Only Connecticut, Nebraska, and Vermont do not have named virtual schools. In those states that do, the schools are run by the local district, state district, or by an education management organization, such as K12, a for- profit company that develops educational curriculum and learning platforms (NCES, 2013: Wikipedia, 2013).

History of virtual education

The first virtual school, Laurel Springs, opened its doors in California in 1991. By 1994 it offered a complete high school curriculum online. Three years later two other public virtual schools emerged, Florida Virtual School and Utah's Electronic High School. Florida Virtual School was originally opened to provide Advanced Placement courses to students throughout the state in districts that did not offer or could not afford an Advanced Placement program (Johnston & Barbour, 2013). In the late 1990s, Massachusetts Virtual High School also was created. The model for the Massachusetts' virtual high school was unique. Districts were given an allotment of slots for their students in exchange for the use of one teacher to teach an online course (Archambault & Kennedy, 2012).

All states offer some form of virtual school today. Most fall into one of four categories--blended programs, single district programs, multiple district programs, or state-wide programs--while a few are virtual charter schools (Spitler, Repetto, & Cavanaugh, 2013). Virtual charter students only make up 2% of the charter school population (Brady, Umpstead & Eckes, 2010). *Keeping Pace with K-12 Online and Blended Schools* notes that 31 states and the District of Columbia had fully online virtual schools operating in the 2012-2013 school year (Watson *et al*., 2013). By 2009-2010, 74% of students in the United States taking virtual courses were high school students. Those taking virtual courses represented approximately 1.3 million high school students in about 53% of high school districts that school year. This number is significantly higher than in the 2004-2005 school year, when 310,000 high school students enrolled in virtual courses, or in the 2002-2003 school year, when only 222,000 did. Virtual schools also have a geographical presence, with a higher percentage of high schools offering students virtual courses in the Southeast (76%) than in the Northeast (38%) (NCES, 2013).

During the 2011-2012 school year the growth of virtual K-12 schools slowed in states where relatively large populations of students in virtual school already existed and in states with longer histories of virtual k-12 schools. However, states with new programs and smaller student populations continue to grow exponentially. For example, the Georgia virtual school population grew 112% from the 2010-2011 school year to the 2011-2012 school year (Watson *et al*., 2013). Some states, such as Michigan, New Mexico, Alabama, and Idaho, all have instituted a mandate that now requires students to take at least one course in the virtual environment prior to graduation (Archambault & Kennedy, 2012).

Virtual school appeal

The popularity of virtual school has increased due to the flexible hours, increased educational opportunities, and increased access to educational resources (Brady *et al*., 2010; Spitler *et al*., 2013). Many student athletes, performers, and entertainers are attracted to these schools due to the accommodating nature of the programs. Instruction can be tailored to meet individual student needs and learning styles (Barbour & Reeves, 2009). Many virtual programs base credit for a course on mastery instead of seat time (Archambault & Kennedy, 2012). As Brady *et al*. (2010) note,

> This unlimited course delivery approached is praised for offering personalized , self-paced instruction, a greater variety of curriculum choices, and "master based learning," which provides for wither an acceleration or deceleration of learning to match the individual student's learning styles and specific educational needs. In addition, cyber charter schools' online format has students using high-tech, "twenty-first century skills" every day.

Students choosing virtual schools in Florida for access to Advanced Placement courses are achieving great success. Florida Virtual School students earned qualifying scores (scores of 3, 4 or 5 on the AP exam) at a higher percentage, at 51%, than their brick-and-mortar peers, at

45%, in 2009. The national average of all students making qualifying scores was 56%. In 2010, 53% of Florida virtual school students earned qualifying scores and 55% in 2011. Only 43% of brick-and- mortar students in Florida schools earned qualifying marks in 2010 and 2011. When surveyed, those students who opted for the virtual model of instruction over traditional brick-and-mortar option explained that they liked that the course was accessible at any time and that they could work at their own pace (Johnston & Barbour, 2013).

One myth about virtual schools is that these programs cater to gifted or accelerated students such as these, when in fact many programs offer courses to students of all levels (Rose & Blomeyer, 2007). Indeed, at-risk students who have not been successful in traditional brick-and-mortar schools are looking to virtual schools as an alternative model and last resort. These schools offer students opportunities to recover credits from previous schools and a more flexible schedule to earn additional credits (Rose & Blomeyer, 2007). Students with behavior issues in traditional brick-and-mortar schools are also finding virtual schooling a practical alternative (Bardour & Reeves, 2009).

Other reasons parents and students are choosing virtual school over traditional brick-and-mortar schools include but are not limited to professional sports or entertainment obligations, illness, pregnancy, fears of school violence, or religious beliefs. . Homeschooled students have also begun to use virtual schools as a means to provide curriculum that parents may not know (Bardour & Reeves, 2009).

Roblyer (2006) reports on three different success stories: Gisselle, who was able to finish school early to travel with a ballet company; Leslie, who was able to finish high school on time after missing half a school year due to pregnancy; and Sidney, who needed credit recovery to graduate on time because his mother felt that a night school option was unsafe. These students

were able to complete high school on time or early as a result of the virtual school offerings in their states.

Differences between brick-and-mortar and virtual schools

There are more differences between brick-and-mortar schools than just physical location. Virtual schools are not just limited to the 6- to 7-hour day traditional schools offer; curriculum in these schools is accessible by students 24 hours a day, seven days a week. Virtual classrooms do not have the same capacity limitations that brick-and-mortar schools experience: the fire marshal is not going to close the school for having too many students in a classroom. Students do not have to live within a certain radius of the virtual school to attend either because many virtual schools are statewide. Virtual private schools often cost less than the brick-and-mortar equivalent (McFarlane, 2011). There are, however, some differences that might be seen as negative. Students do not have much if any face to face contact with teachers. Feedback on assignments some may argue might not be as timely or as effective as in traditional brick-and-mortar schools.

Virtual school may not be the best choice for all students. Reid, Aqui, and Putney (2009) found that administrators surveyed believe that successful virtual school students need to be independent and highly motivated and have good family support and an ability to manage time. The survey reveals that students who left the virtual school feel isolated and have unrealistic expectations of the realities of the program. More than 60% of the educators surveyed in the Piccano and Seaman (2007) study report that students need to be more self-disciplined to complete online courses than they do in face-to-face courses. The educators surveyed also note that online offerings allowed smaller districts to offer a larger variety of courses (Piccano & Seaman, 2007).

Summary

The accountability movement has evolved for the last 50 years, from a movement in equality of education to a movement concerned primarily with student and school achievement. The measures put into place by the Elementary and Secondary Education Act provided Title 1 funds meant to go to schools with high rates of poverty. The report from the 1980s shows that districts were not using these funds appropriately nor were the funds making the difference legislators had anticipated (McGuinn, 2006). The Reagan Administration, following *The Nation at Risk* report, urged states to hold students and schools accountable for achievement. This birthed the first of the high-stakes achievement tests for the nation. Not all states willingly began looking at achievement, and in 2001 NCLB forced those states who were not already judging student and school achievement to do so.

This was not the first time students were judged by achievement. Students have been receiving grades for academic achievement since the birth of the one-room school house but never before has there been such discussion of what actually makes up a grade or whether grades are an accurate depiction of student knowledge. Research shows that teachers take into account ability, behavior, and effort when awarding grades. Whether due to this or other reasons, grades point averages have steadily increased over the last 40 years. While grade point averages have increased, the average ACT/SAT score has stayed relatively the same. More and more students are applying to college than ever before, but many of these students are unprepared and have to take remediation courses in English and mathematics before starting core classes. Graduate and professional schools now have to look at other indicators for admissions due to the fact that post-secondary grades also suffer from inflation.

Many states now have policies that make high-stakes exam grades part of the students' overall course grade. This in turn causes teachers and schools to make passing the test a top priority. It seems as long as high-stakes testing continues to be a top priority for educators, curriculum narrowing will occur. At some point educators and legislators will have to address whether the tests cause more harm than good. Warren and Grodsky (2009) indicate that high-stakes exams do not benefit those who pass them and do not prepare students better for the labor market. Berliner (2009) shows that SAT scores are better in states that do not require high-stakes tests than in those that do, but other research indicates that SAT scores improve among particular student populations when high-stakes tests are in place. States manipulate passing scores in order to produce the desired pass rate (Yaffee, 2009).

High-stakes tests have caused some teachers to change their pedagogy, shifting to the use of lecture and rote memorization (Au, 2009). High-stakes tests also cause teachers to leave out or vastly modify curriculum and instruction (Nichols & Berliner, 2006). Martone and Sireci (2010) state that if gaps in curriculum alignment exist, measures should be taken to correct the problem.

While ESEA and NCLB assume that poverty rates are linked to low achievement on high-stakes tests instead of a lack of well-paying jobs, research shows that is not the case (Anyon & Greene, 2007).

Virtual education is becoming a popular choice for parents and students (Brady *et al*., 2010; Spitler *et al*., 2013). These programs offer flexibility, myriad course options, and opportunities for individualized instruction not often available in traditional brick-and-mortar settings (McFarlane, 2011). Students attend these schools for a variety of reasons. Virtual

schools offer opportunities for all students, from advanced learners all the way to at-risk learners. However, virtual schools may not be the best fit for all students (Piccano & Seaman, 2007).

CHAPTER THREE: METHODOLGY

Introduction

Teachers are under immense pressure to increase course pass rates and increase EOCT scores. In the era of NCLB accountability, states and school districts must show that students are mastering the curriculum. This mastery is proven by the students' performance on high-stakes tests. In many states, including Georgia, students must pass a battery of high-stakes criterion referenced tests in order to receive a high school diploma. Many states have also put achievement tests into place that determine whether students are promoted to the next grade. Student performance has always been a priority for teachers, but never with the current intensity. Teachers are now under immense pressure from the federal government to increase test scores and therefore they direct instruction to prepare students for the test, resulting in a narrowing of the curriculum. The new education initiatives of the Obama Administration are trying to directly link teacher pay with student achievement scores, as part of RTTI (American Recovery and Reinvestment Act, 2009). With student test scored directly linked to teacher pay increases, focus in classrooms will be on standards that will be tested more than ever before.

This study will attempt to determine whether there is a relationship between student course scores and scores on the accompanying End of Course tests. A public state virtual high school in Georgia has been chosen for this study. Approximately 4,000 high school students are enrolled at this virtual school. The focus of this study is to determine if end of course grades are an accurate reflection of the content knowledge as indicated by measurement in state's End of Course tests.

Design

This study will be a correlational study. A correlational design is the best method for determining the relationship between the variables under investigation. Correlational research

attempts to determine the strength of the relationship between two variables, in this case the end of course grades and the EOCT scores. This design is quantitative in nature as this study seeks to take numerical data, perform statistical analysis, and to predict trends (Gall, Gall, & Borg, 2007). Student end of course grades and EOCT scores will be gathered in order to determine if a relationship exists between the scores. End of course grades (percentages) should be predictive of the students' mastery of curriculum. EOCT scores should reflect the student mastery of curriculum.

Variables

End of course grades

EOCT scores

Question and Hypothesis

Research Question 1: Is there a relationship between end of course grades and EOCT scores in Algebra, Geometry, Biology, Physical Science, Ninth Grade Literature, and U.S. History courses?

$H_0$1 There is no statistically significant correlation between student end of course grades (shown as percentages) and student EOCT scores in Ninth Grade Literature.

$H_0$2 There is no statistically significant correlation between student end of course grades (shown as percentages) and student EOCT scores in Algebra.

$H_0$3 There is no statistically significant correlation between student end of course grades (shown as percentages) and student EOCT scores in Geometry.

$H_0$4 There is no statistically significant correlation between student end of course grades (shown as percentages) and student EOCT scores in Biology.

$H_0$5 There is no statistically significant correlation between student end of course grades (shown as percentages) and student EOCT scores in Physical Science.

$H_0$6 There is no statistically significant correlation between student end of course grades (shown as percentages) and student EOCT scores in U.S. History.

Research Question 2: Is there a relationship between end of course grades and EOCT scores for students with disabilities?

$H_0$7 There is no statistically significant correlation between student end of course grades (shown as percentages) and student End of Course scores for regular education students and students with disabilities.

Research Question 3: Is there an association between pass rates (earning a 70% or higher) in the course and pass rates (earning a 70% or higher) on the End of Course tests required for that course?

$H_0$8 There is no association in passing (earning a 70% or higher) in the course and passing (earning a 70% or higher) on the End of Course Test required for that course.

Participants

Approximately 2,300 high school students' end of course grades and EOCT scores will be used for this study. The students at this school reside all over the State of Georgia. These students complete all of their courses in a virtual environment but are required to take the End of Course tests at assigned testing locations around the state.

Setting

A state public chartered virtual high school in Georgia has been chosen for this study. The central office for this school is located in Atlanta. The students take the required End of Course tests at various testing sites across the state. Once materials are delivered to the school

from the state, the materials are transported in locked boxes by certified staff members to the various locations and then returned to the school's central office. The school's central office in turn delivers the testing materials back to the state for scoring.

Instrumentation

This study will utilize the EOCT scores in Algebra, Geometry, Biology, Physical Science, Ninth Grade Literature, and U.S. History provided from the state. The tests are designed to measure student knowledge in each of the core areas, and the tests have been normed for the intended population. The instrument was developed by the State of Georgia, which has deemed the instrument reliable. The Georgia Department of Education reported the reliability levels of all the End of Course tests to range between .86 and .94 (Cox, 2006). The Georgia Department of Education and the federal government have also deemed these tests valid, reliable, and objective for the use of determining the state's AYP status. The Georgia Department of Education defines reliability as the "extent to which an examinee's performance is consistent over time" and validity as "how well the items measure what they are intended to measure and the extent to which the inferences drawn from the scores are supported" (Cox, 2006).

Raw teacher-assigned numerical end of course grades were utilized for this study. Final end of course grades consist of a calculation of the teacher-assigned grade (either 80% or 85% of the total grade, depending on the year the student entered high school) and the EOCT score (the remaining 15% or 20%, depending on the year the student entered high school). Course teachers assign numerical end of course grades based upon student performance in the course. Students who earn a combined score of 70% or higher receive one unit of credit for the course. Students

who earn a combined score of 69% or lower must retake the course. For this research a 100-90 equates to an A, 89-80 equates to a B, 79-70 equates to a C, and 69 or below equates to an F.

Procedures

Approval to conduct this study was sought from the Liberty University Institutional Review Board (IRB). Once IRB approval was received, end of course grade and EOCT data were obtained from the 2012-2013 school year from a virtual high school in the State of Georgia. Course scores from Algebra, Geometry, Biology, Physical Science, Ninth Grade Literature, and U.S. History were gathered. American Literature and Economics scores were not used due to the limited number of students who took these courses during the 2012-13 school year. The data used to complete this study were obtained from the virtual high school's central office. The school accesses current course grades weekly from an online data base. The final grades issued by the teachers, prior to the EOCT scores being calculated into the average, were used for the study. Data analyses were performed on each individual course and correlating test.

Data Analysis

A correlation study was performed using Pearson Product-Moment Correlation tests to determine if there is a statistically significant relationship between numerical end of course grades as awarded by the course teacher and scores on the End of Course tests required for that course. The Pearson Product-Moment Correlation test provided the researcher with a correlation coefficient, which allowed the researcher to determine the strength of the relationship between the numerical end of course grade earned in a course and the numerical score earned on the accompanying End of Course Test. The analyses were repeated separately for regular education students and for students with disabilities in order to examine the relationship for each group individually. McNemar's chi square test was conducted to determine if there is a statistically

significant association between passing the course and passing the End Of Course Test associated with the course. McNemar's chi squared test allowed the researcher to evaluate the association between two separate variables for the same individuals. All statistical tests were conducted at the .05 level of significance.

CHAPTER FOUR: FINDINGS

Introduction

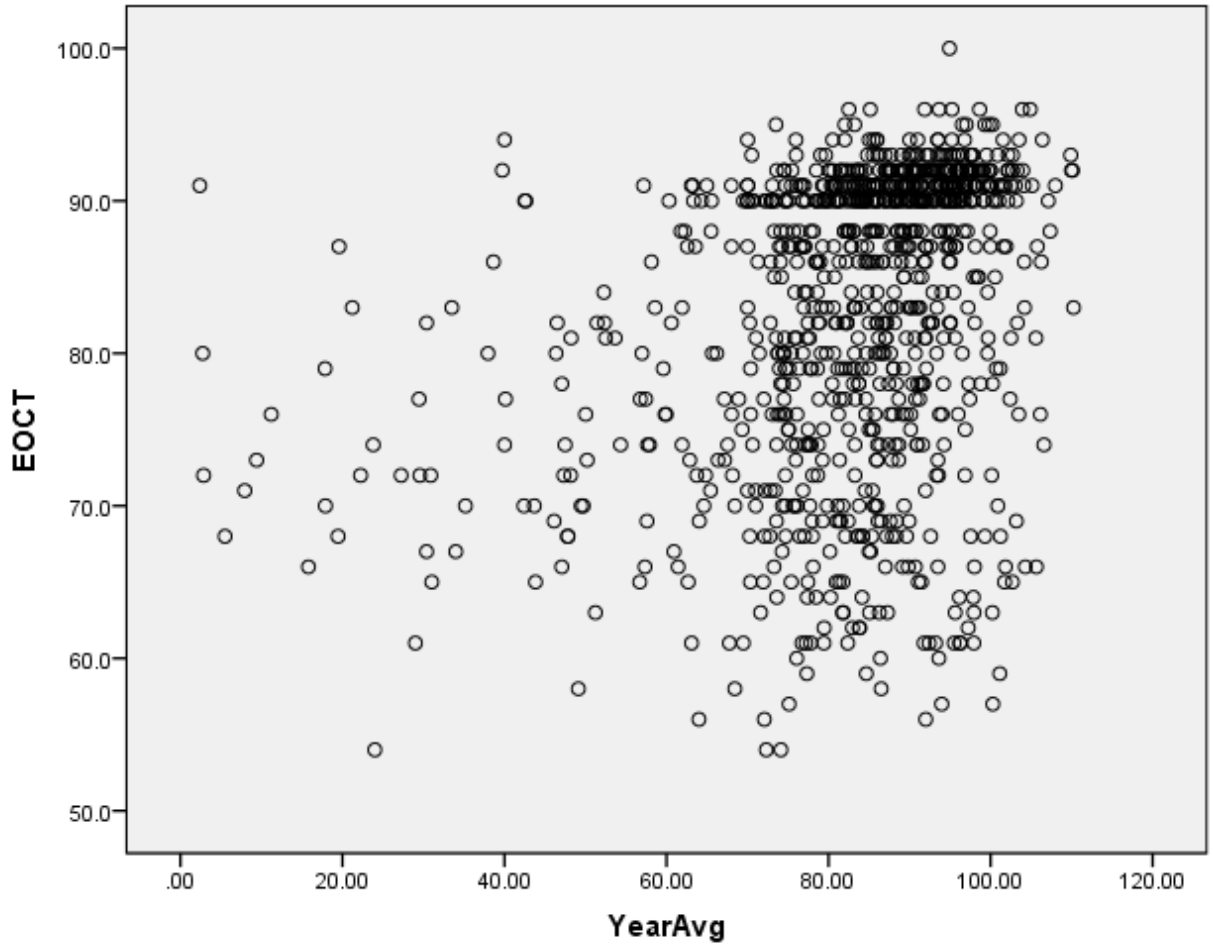
The purpose of this study was to investigate the relationship between end of course grades and the EOCT scores associated with courses in Ninth Grade Literature, Algebra, Geometry, Biology, Physical Science, and U.S. History at one virtual school in Georgia. The study also investigated these relationships for students with disabilities. The study sought to determine if there is an association between passing a course (earning a 70% or higher) and passing (earning a 70% or higher) the End of Course Test for each course.

Research Question Findings: Pearson Product-Moment Correlation

Overall Ninth Grade Literature

The average Ninth Grade Literature student had an EOCT score of 82.56 (SD = 9.87), and a course average of 82.66 (SD = 16.33). A non-significant result from Kolmogorov-Smirnov's test (p = .000) shows that the distribution of EOCT scores and end of course grades is not normal. Figure 1 shows a scatterplot of the EOCT scores and course points for each student, showing that the assumptions of linearity and homoscedasticity are not tenable.

Figure 1



Pearson product-moment correlation was used to evaluate the null hypothesis show that there is no relationship between students' Ninth Grade Literature EOCT scores and their end of course grades (n = 968). There was evidence to reject the null hypothesis and conclude that there is a small to moderate positive correlation between EOCT scores and end of course grades, r = .31, p = .000 (Cohen, 1988). It appears that higher EOCT scores are moderately related to higher overall end of course grades.

General Education verses Students with Disabilities for Ninth Grade Literature

Separate Pearson analyses were conducted for students with disabilities and regular education students. The results show that the average Ninth Grade Literature regular education student had an EOCT score of 83.51 (SD = 9.34), and a course average of 82.84 (SD = 16.09). Pearson product-moment correlation was used to evaluate the null hypothesis that no relationship between the Ninth Grade Literature EOCT scores of regular education students and their end of course grades (n = 864). There was significant evidence to reject the null hypothesis and conclude that there is a moderate positive correlation between EOCT scores and end of course grades, r = .36, p = .000 (Cohen, 1988).

The average Ninth Grade Literature special education student had an EOCT score of 74.67 (SD = 10.65), and a course average of 81.12 (SD = 18.22). However, there was not significant evidence to reject the null hypothesis. The data show that there is little association, r = .02, p = .83 between the Ninth Grade Literature EOCT scores of students with disabilities and their end of course grades (n = 104).

Table 1

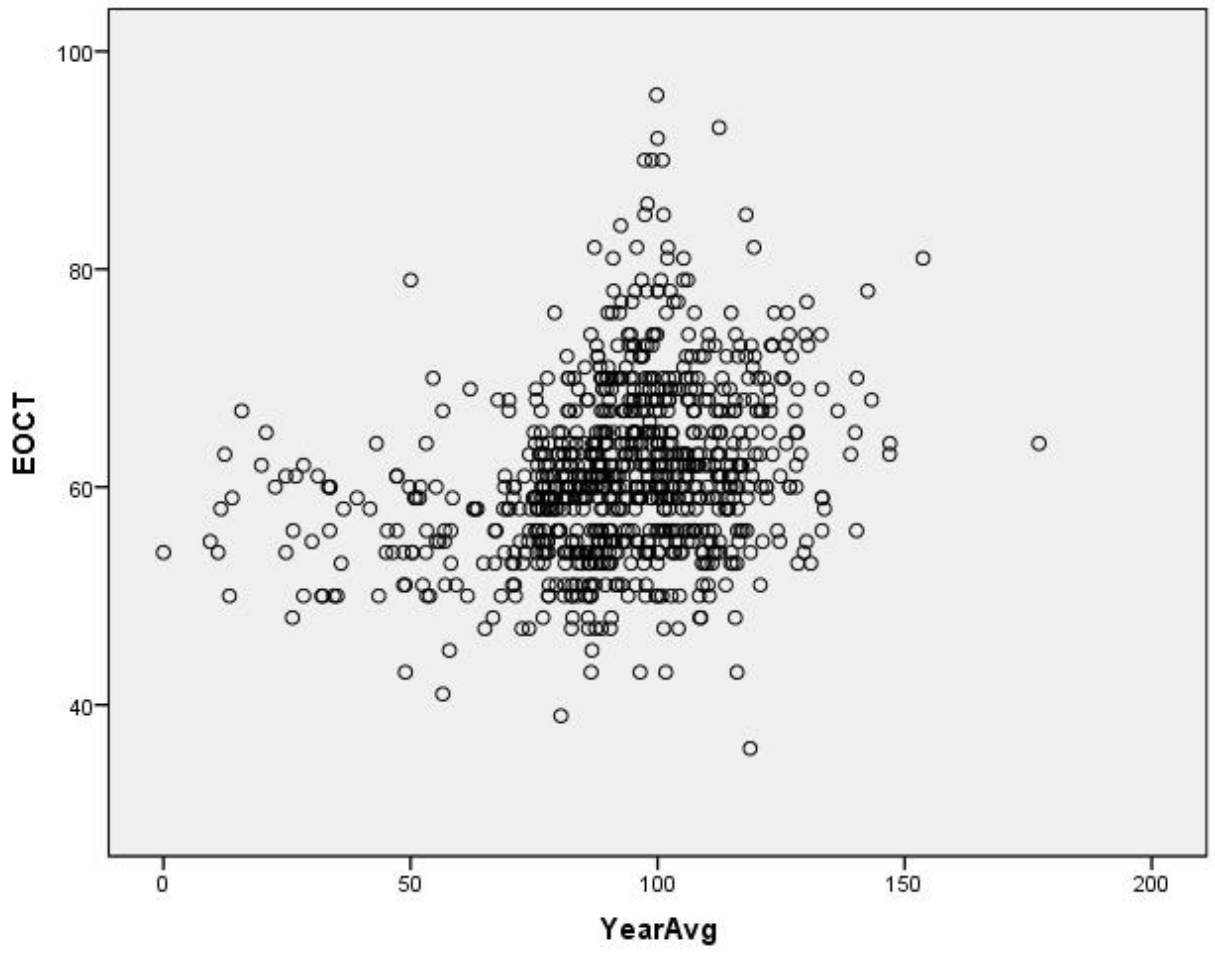
Pearson product-moment correlation on Ninth Grade Literature

Ninth Grade Literature	*N*	*r*	*p*		Mean	SD
Whole Group	968	.31	.000	Course	82.66	16.33
				EOCT	82.56	9.87
Special Education	104	.02	.83	Course	81.12	18.22
				EOCT	74.67	10.65

Overall Algebra

The average Algebra student had an EOCT score of 61.26 (SD = 7.97), and a course average of 92.26 (SD = 22.39). A non-significant result from Kolmogorov-Smirnov's test (p = .000) shows that the distribution of both EOCT scores and end of course grades is not normal. Figure 2 shows a scatterplot of the EOCT scores and course points for each student, showing that the assumptions of linearity and homoscedasticity are not tenable.

Figure 2



Pearson product-moment correlation was used to evaluate the null hypothesis that there is no relationship between the Algebra EOCT scores of students and their end of course grades (n =

880). There was evidence to reject the null hypothesis and conclude that there is a weak, positive correlation between EOCT scores and end of course grades, $r = .28$, $p = .000$ (Cohen, 1988). It appears that higher EOCT scores are moderately related to higher overall end of course grades.

General Education verses Students with Disabilities for Algebra

Separate tests were run, one for student with disabilities and one regular education students. The average regular education Algebra student had a EOCT score of 61.80 (SD = 7.97), and a course average of 92.95 (SD = 22.15) and the average special education Algebra student had an EOCT score of 57.53 (SD = 7.04), and a course average of 88.36 (SD = 22.06). Pearson product-moment correlation was used to evaluate the null hypothesis that there is no relationship between the Algebra EOCT scores of regular education students and their end of course grades ($n = 768$) and to evaluate Algebra EOCT scores of students with disabilities and their end of course grades ($n = 112$). There was evidence to reject the null hypothesis and conclude that there is a weak, positive correlation between EOCT scores and end of course grades, $r = .28$, $p = .000$ for regular education students and there was significant evidence to reject the null hypothesis and conclude that there is a weak, positive correlation between EOCT scores and end of course grades, $r = .28$, $p = .003$ for students with disabilities (Cohen, 1988).

Table 2

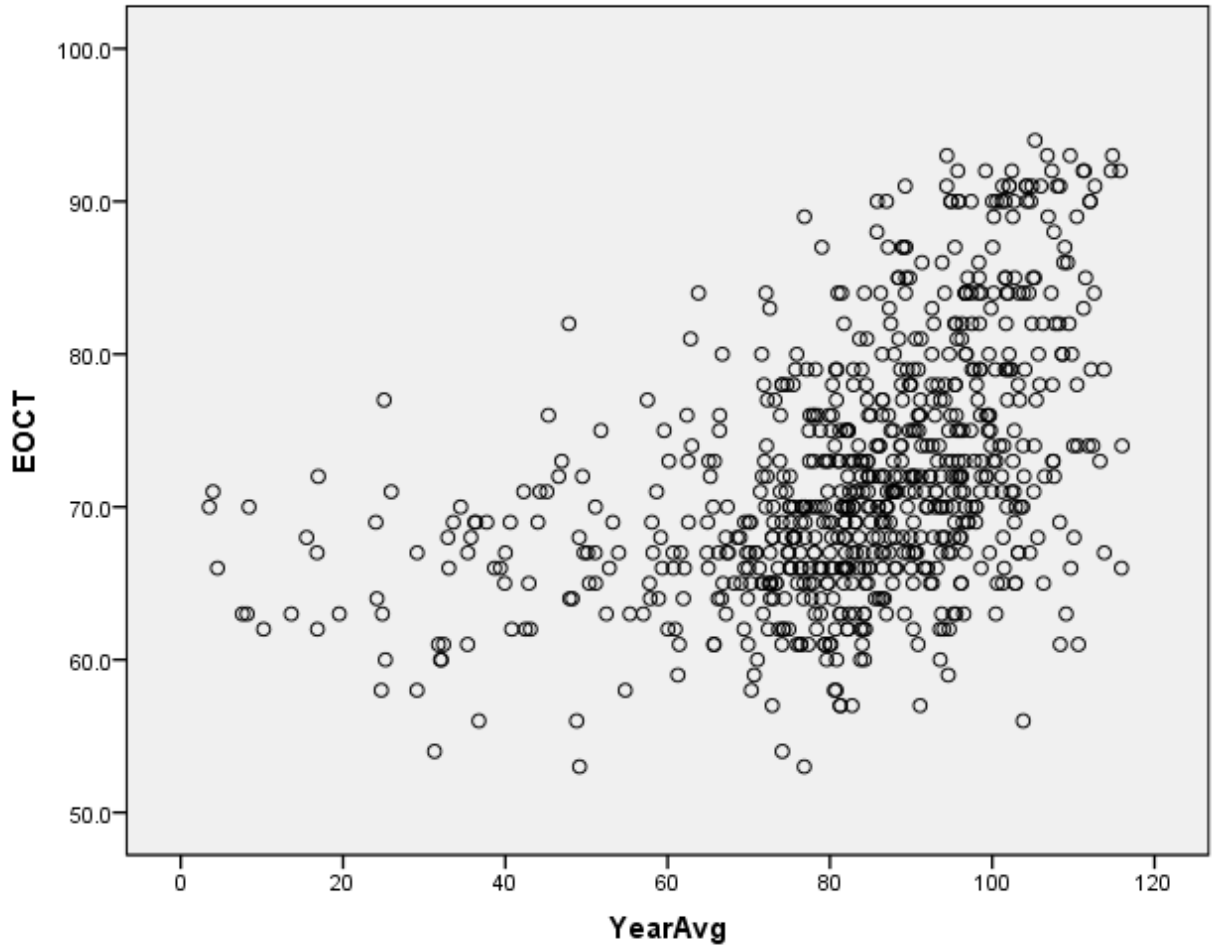
Pearson product-moment correlation on Algebra

Algebra	*N*	*r*	*p*		Mean	SD
Whole Group	880	.28	.000	Course	92.26	22.39
				EOCT	61.26	7.97
Special Education	112	.28	.003	Course	88.36	22.06
				EOCT	57.53	7.04

Overall Geometry

The average Geometry student had an EOCT score of 72.11 (SD = 8.27), and a course average of 82.51 (SD = 20.12). A non-significant result from Kolmogorov-Smirnov's test (p = .000) shows that the distribution of both EOCT scores and end of course grades is not normal. Figure 3 shows a scatterplot of the EOCT scores and course points for each student, showing that the assumptions of linearity and homoscedasticity are not tenable.

Figure 3



Pearson product-moment correlation was used to evaluate the null hypothesis that there is no relationship between the Geometry EOCT scores of students and their end of course grades (n = 764). There was evidence to reject the null hypothesis and conclude that there is a moderate, positive correlation between EOCT scores and end of course grades, r = .46, p = .000 (Cohen, 1988). It appears that higher EOCT scores are moderately related to higher overall end of course grades.

General Education verses Students with Disabilities for Geometry

Separate analyses were conducted, one for students with disabilities and one without students with disabilities. The average Geometry regular education student had an EOCT score of 72.33 (SD = 8.29), and a course average of 83.08 (SD = 19.54) and the special education student had an EOCT score of 67.90 (SD = 6.59), and a course average of 71.64 (SD = 27.23). Pearson product-moment correlation was used to evaluate the null hypothesis that there is no relationship between the Geometry EOCT scores of regular education students and their end of course grades (n = 726) and students with disabilities end of course grades (n = 38). There was evidence to reject the null hypothesis and conclude that there is a moderate, positive correlation between EOCT scores and end of course grades, r = .46, p = .000 (Cohen, 1988). However, there was not significant evidence to reject the null hypothesis, r = .29, p = .07 for students with disabilities.

Table 3

Pearson product-moment correlation on Geometry

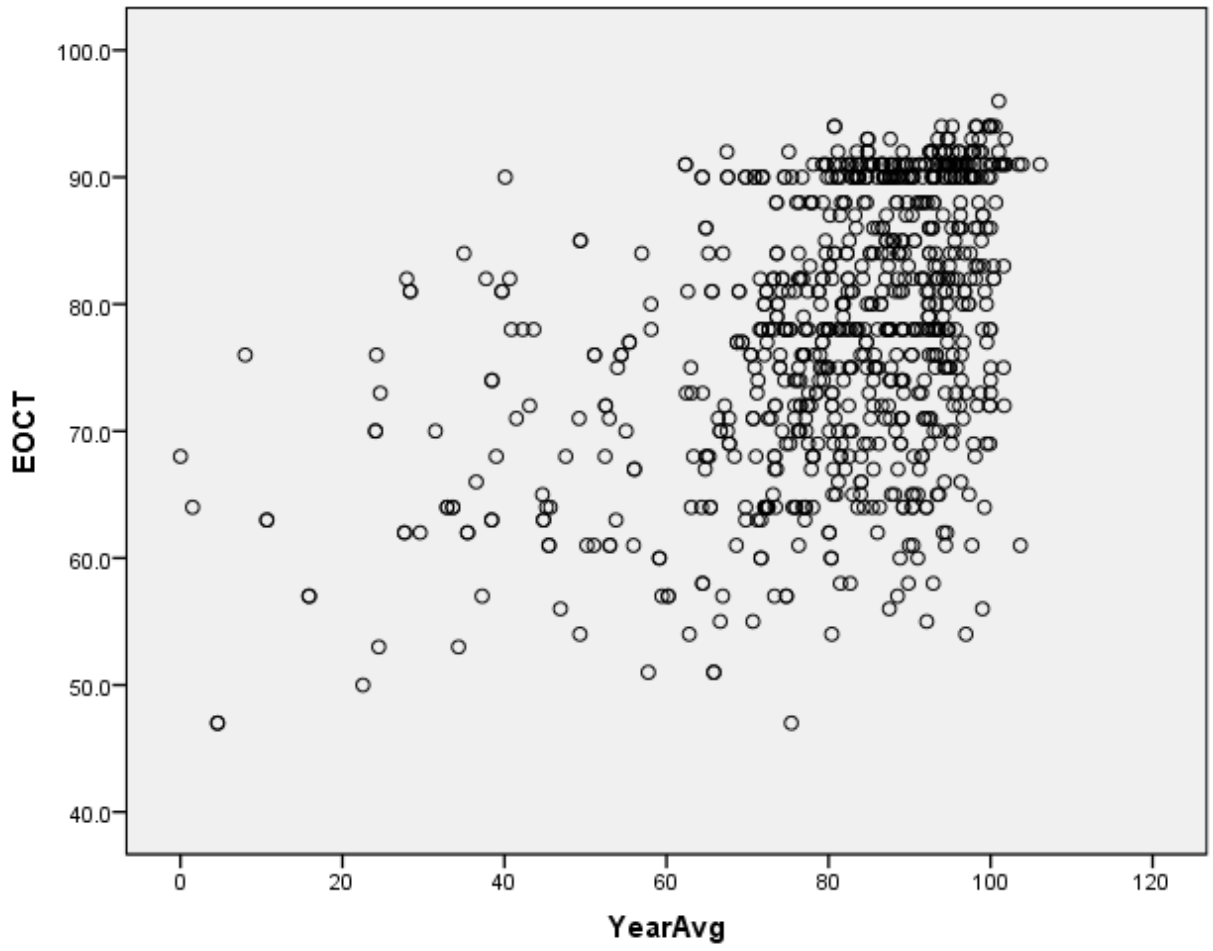
Geometry	*N*	*r*	*p*		Mean	SD
Whole Group	764	.46	.000	Course	82.51	20.12
				EOCT	72.11	8.27
Special Education	38	.29	.07	Course	71.64	27.23
				EOCT	67.90	6.59

Overall Biology

The average Biology student had an EOCT score of 78.04 (SD = 10.53), and a course average of 80.58 (SD = 17.42). A non-significant result from Kolmogorov-Smirnov's test (p = .000) shows that the distribution of EOCT scores and end of course grades is not normal. Figure

4 shows a scatterplot of the EOCT scores and course points for each student, showing that the assumptions of linearity and homoscedasticity are not tenable.

Figure 4



Pearson product-moment correlation was used to evaluate the null hypothesis that there is no relationship between the biology EOCT scores of students and their end of course grades (n = 839). There was evidence to reject the null hypothesis and conclude that there is a moderate, positive correlation between EOCT scores and end of course grades, r = .44, p = .000 (Cohen, 1988). It appears that higher EOCT scores are moderately related to higher overall end of course grades.

General Education verses Students with Disabilities for Biology

Separate analyses were conducted, one for regular education students and students with disabilities. The average Biology regular education student had an EOCT score of 78.13 (SD = 10.36), and a course average of 80.95 (SD = 16.80) and the special education student had an EOCT score of 73.00 (SD = 18.08), and a course average of 58.44 (SD = 33.65). Pearson product-moment correlation was used to evaluate the null hypothesis that there is no relationship between the Biology EOCT scores of regular education students and their end of course grades (n = 825) and the special education student and their end of course grades (n = 14). There was evidence to reject the null hypothesis and conclude that there is a moderate, positive correlation between EOCT scores and end of course grades, r = .41, p = .000, and a moderate to strong positive correlation for students with disabilities, r = .81, p =.000 (Cohen, 1988).

Table 4

Pearson product-moment correlation on Biology

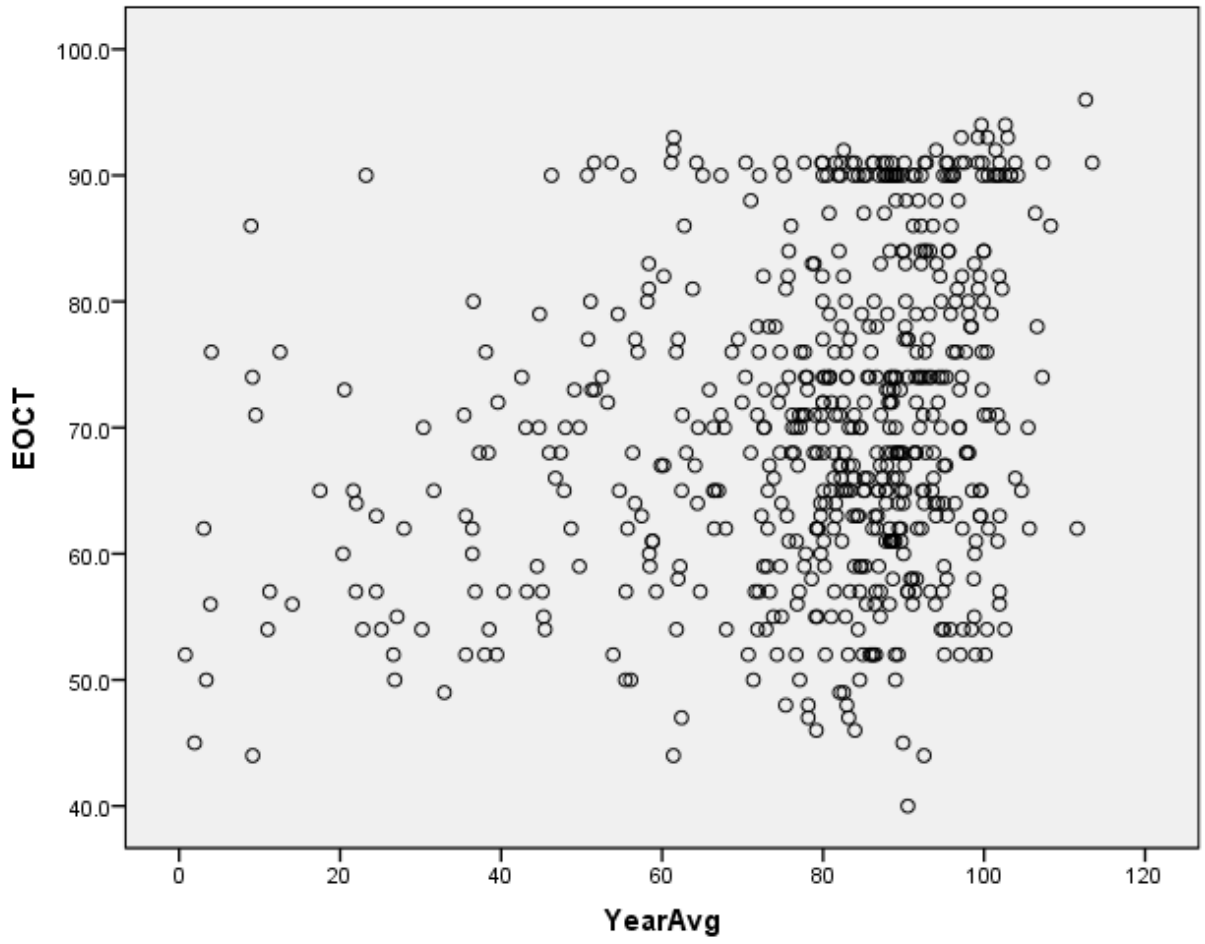
Biology	*N*	*r*	*p*		Mean	SD
Whole Group	839	.44	.000	Course	80.58	17.42
				EOCT	78.04	10.53
Special Education	14	.81	.000	Course	58.44	33.65
				EOCT	73.00	18.08

Overall Physical Science

The average Physical Science student had an EOCT score of 70.41 (SD = 12.53), and a course average of 78.57 (SD = 21.24). A non-significant result from Kolmogorov-Smirnov's test (p = .000) shows that the distribution of EOCT scores and end of course grades is not

normal. Figure 5 shows a scatterplot of the EOCT scores and course points for each student, showing that the assumptions of linearity and homoscedasticity are not tenable.

Figure 5



Pearson product-moment correlation was used to evaluate the null hypothesis that there is no relationship between the Physical Science EOCT scores of students and their end of course grades (n = 610). There was evidence to reject the null hypothesis and conclude that there is a moderate, positive correlation between EOCT scores and end of course grades, r = .25, p = .000 (Cohen, 1988). It appears that higher EOCT scores are moderately related to higher overall end of course grades.

General Education verses Students with Disabilities for Physical Science

Separate analyses were conducted, one for regular education students and students with disabilities. The average Physical Science regular education student had an EOCT score of 71.23 (SD = 12.50), and a course average of 78.73 (SD = 21.18) and the average special education student had an EOCT score of 66.25 (SD = 11.91), and a course average of 77.76 (SD = 21.62). Pearson product-moment correlation was used to evaluate the null hypothesis that there is no relationship between the Physical Science EOCT scores of regular education students and their end of course grades (n = 510) and of students with disabilities and their end of course grades (n = 100). There was evidence to reject the null hypothesis and conclude that there is a weak, positive correlation between EOCT scores and end of course grades, r = .27, p =.000 (Cohen, 1988). However, there was not significant evidence to reject the null hypothesis, r = .17, p = .10 for students with disabilities.

Table 5

Pearson product-moment correlation on Physical Science

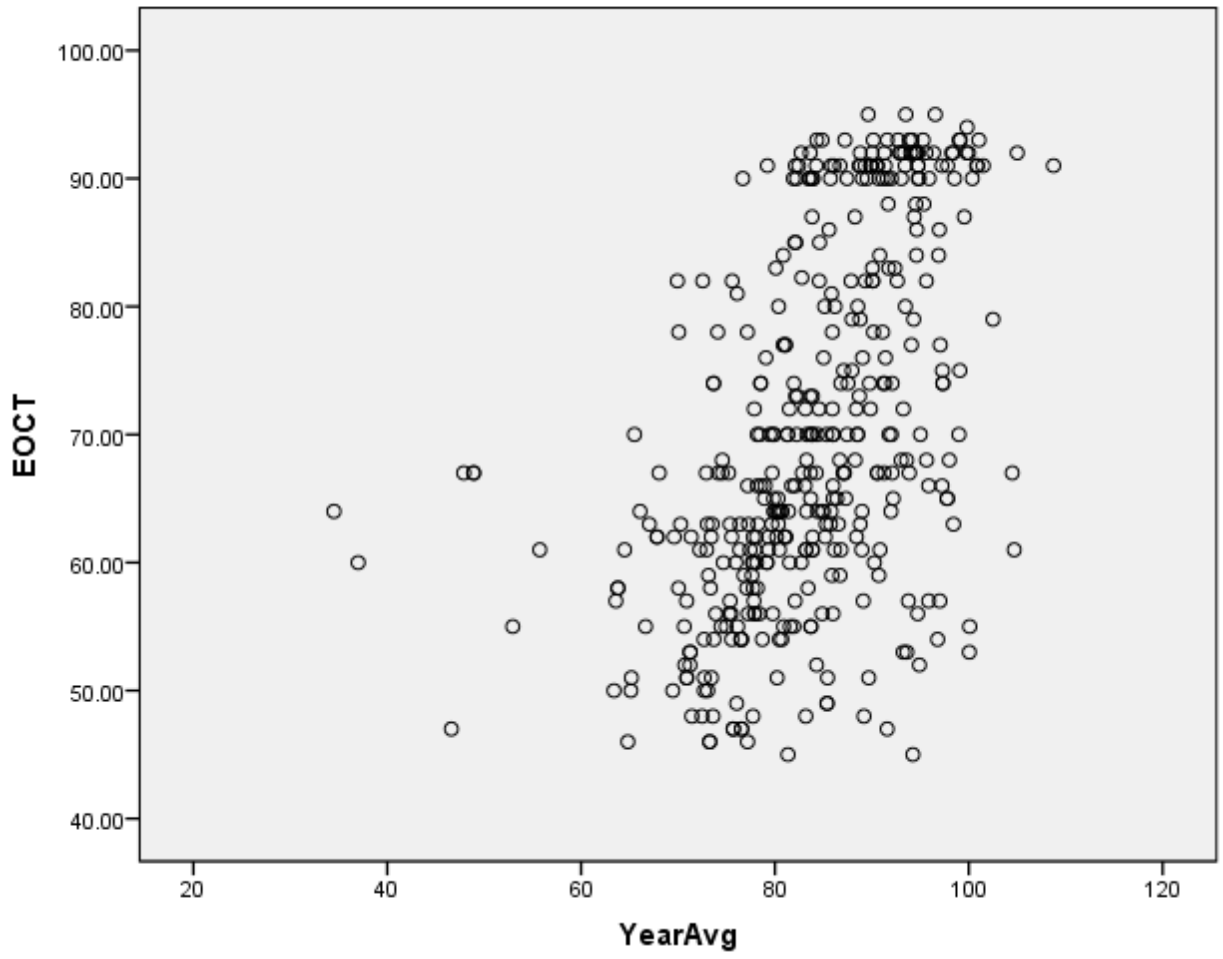
Physical Science	*N*	*r*	*p*		Mean	SD
Whole Group	610	.25	.000	Course	78.57	21.24
				EOCT	70.41	12.53
Special Education	100	.17	.10	Course	77.76	21.62
				EOCT	66.25	11.91

Overall U.S. History

The average U.S. History student had an EOCT score of 70.23 (SD = 13.98), and a course average of 83.79 (SD = 10.20). The analyses from the U.S. History data indicated a significant result from Kolmogorov-Smirnov's test (p = .15) for the course average and indicated

non-significant result for the End of Course test (p = .000). Figure 6 shows a scatterplot of the EOCT scores and course points for each student, showing that the assumptions of linearity and homoscedasticity are not tenable.

Figure 6



Pearson product-moment correlation was used to evaluate the null hypothesis that there is no relationship between the social studies EOCT scores of students and their end of course grades (n = 419). There was evidence to reject the null hypothesis and conclude that there is a moderate to strong, positive correlation between EOCT scores and end of course grades, r = .51,

p = .000 (Cohen, 1988). It appears that higher EOCT scores are moderately related to higher overall end of course grades.

General Education verses Students with Disabilities for U.S. History

Separate analyses were conducted, one for regular education students and students with disabilities. The average US History regular student had an EOCT score of 70.46 (SD = 13.99), and a course average of 83.69 (SD = 9.99) and the average special education student had an EOCT score of 67.14 (SD = 13.71), and a course average of 85.10 (SD = 12.90). Pearson product-moment correlation was used to evaluate the null hypothesis that there is no relationship between the social studies EOCT scores of regular education students and their end of course grades (n = 390) and for students with disabilities and their end of course grades (n = 29). There was evidence to reject the null hypothesis and conclude that there is a moderate to strong, positive correlation between EOCT scores and end of course grades, r = .51, p = .000 for general education students, as well as a moderate correlation for students with disabilities, r = .48, p = .008 (Cohen, 1988).

Table 6

Pearson product-moment correlation on U.S. History

US History	*N*	*R*	*p*		Mean	SD
Whole Group	419	.51	.000	Course	83.79	10.20
				EOCT	70.23	13.98
Special Education	29	.48	.008	Course	85.10	12.90
				EOCT	67.14	13.71

Research Question Findings: McNemar's Chi Square

A McNemar's chi squared test on paired proportions was conducted to evaluate the relationship between passing a course and passing the EOCT for Ninth Grade Literature, Algebra, Geometry, Biology, Physical Science, and U.S. History classes. In Ninth Grade Literature most students who had failing end of course grades (77.1%) had passing EOCT scores. However, most students who had passing end of course grades (87.4%) also had passing EOCT scores. The results of the McNemar test were not significant, χ^2= 1.136, df = 1 *p*= .28 (see table 9 below). There was not significant evidence to reject the null hypothesis that there is no relationship between passing the course and passing the End of Course Test in Ninth Grade Literature. There is no statistically significant relationship between passing the course and a passing EOCT score in Ninth Grade Literature.

Table 7

McNemar's Chi Square on Ninth Grade Literature Course Pass Rates and EOCT Pass Rates

EOCT		Fail	Pass
Course	Fail	27	91
	Pass	107	743

Ninth Grade Literature χ^2=1.136, df=1, p=.28 N = 968

Students in Algebra who had failing end of course grades (97.8%) also had failing EOCT scores. However, most students who had passing end of course grades (89%) also had failing EOCT scores. The results of the McNemar test were significant, χ^2 742.03, df=1, *p* = .000 (see Table 10 below). There was significant evidence to reject the null hypothesis and conclude that there is a relationship between passing a course and passing the End of Course Test in Algebra.

Table 8

McNemar's Chi Square on Algebra Course Pass Rates and EOCT Pass Rates

EOCT		Fail	Pass
Course	Fail	91	2
	Pass	750	128

Algebra χ^2=742.03, df=1, p=.000 N = 881

Students in Geometry who had failing end of course grades (74.4%) also had failing EOCT scores. However, most students who had passing end of course grades (63.4%) had passing EOCT scores. The results of the McNemar test were significant, χ^2= 144.96, df= 1 *p* = .000 (see Table 11 below). There was evidence to reject the null hypothesis that there is no relationship between passing the course and passing the End of Course Test in Geometry.

Table 9

McNemar's Chi Square on Geometry Course Pass Rates and EOCT Pass Rates

EOCT		Fail	Pass
Course	Fail	99	34
	Pass	231	400

Geometry χ^2=144.96, df=1, p=.000 N= 764

Students in Biology who had failing end of course grades (50.3%) had failing EOCT scores. However, most students who had passing end of course grades (83%) also had passing EOCT scores. The results of the test were significant, χ^2 = 11.19, df = 1, *p* = .000 (see table 12

below). There was evidence to reject the null hypothesis that there is no relationship between passing the course and passing the End of Course Test in Biology.

Table 10

McNemar's Chi Square on Biology Course Pass Rates and EOCT Pass Rates

EOCT		Fail	Pass
Course	Fail	72	71
	Pass	118	578

Biology χ^2=11.19, df=1, p=.000 N = 839

Physical Science students who had failing end of course grades (59.0%) had failing EOCT scores. However, most students who had passing end of course grades (53.7%) also had passing EOCT scores. The results of the test were significant, χ^2= 97.80, df=1, p = .000 (see table 13 below). There was evidence to reject the null hypothesis that there is no relationship between passing the course and passing the End of Course Test in Physical Science.

Table 11

McNemar's Chi Square on Physical Science Course Pass Rates and EOCT Pass Rates

EOCT		Fail	Pass
Course	Fail	79	55
	Pass	220	255

Physical Science χ^2=97.80, df=1, p=.000 N = 610

U.S. History students who had failing end of course grades (92.3%) also had failing EOCT scores. However, most students who had passing end of course grades (53.5%) also had

failing EOCT scores. The results of the test were significant, $\chi^2 = 213.11$, $p = .000$ (see table 14 below). There was significant evidence to reject the null hypothesis that there is no relationship between passing the course and passing the End of Course Test in US History.

Table 12

McNemar's Chi Square on US History Course Pass Rates and EOCT Pass Rates

EOCT		Fail	Pass
Course	Fail	24	2
	Pass	201	192

US History χ^2=193.12, df=1, p=.000 N = 419

Summary

The results from the Pearson product-moment correlation analyses indicate that the researcher should reject the null hypothesis due to the fact that there is a weak to moderate positive relationship between passing a course and passing the End of Course tests associated for each course. When separate tests were conducted for regular education students and students with disabilities, the Pearson product-moment correlation test indicates that the null hypothesis could not be rejected for Ninth Grade Literature, Geometry, and Physical Science for students with disabilities. The null hypothesis could be rejected for all other subjects, as they indicated weak to moderate positive correlation between passing a course and passing the End of Course Test for students with disabilities. The analyses indicate that the researcher should reject the null hypothesis for regular education students for each test. The McNemar's chi squared test on paired proportions indicated that there was an association between passing a course and passing the End of Course Test for the course in all subjects tested except for Ninth Grade Literature.

CHAPTER FIVE: DISCUSSION

Summary of the Research Results

The purpose of this study was to discover the relationship between passing a course and passing the End of Course Test associated with each course. Understanding the relationship will allow the virtual school studied to better instruct students and to make policies that ensure the success of the students' academic achievement on standardized tests. The academic achievement data is one of the main items used to calculate the CCRPI score for schools in Georgia. Many students are failing to meet the standards on End of Course tests but passing the course associated with the test.

Pearson product-moment correlation analyses were conducted to examine the relationship between scores earned in a course and EOCT scores for that course. The results from the tests indicate that there is a positive relationship between end of course grades and EOCT scores associated for all courses examined. The results showed that higher EOCT scores are associated with higher overall end of course grades. The positive relationship was stronger with the Geometry, Biology, and U.S. History grades and scores. Ninth Grade Literature and Physical Science had moderate positive relationships between grades and scores, while Algebra had the weakest positive relationship.

When the results for students in regular education were compared to the results for students with disabilities, Pearson product-moment correlation tests did not have the same results. The results for students with disabilities showed that the relationship between end of course grades and EOCT scores was not significant for Ninth Grade Literature, Geometry, and Physical Science. The McNemar chi squared test on paired proportion was conducted on each set of course data. The results from the tests showed that there is a statistically significant

association between passing a course and passing the End of Course Test for the course in all subjects tested except for Ninth Grade Literature.

Discussion of Research Results

The first research question for this study was: is there a relationship between end of course grades and EOCT scores in Ninth Grade Literature, Algebra, Geometry, Biology, Physical Science, and U.S. History courses? All of the results from the Pearson's product-moment correlation tests indicate a statistically significant correlation between end of course grades and EOCT scores of various strengths. U.S. History, Biology, and Geometry results show moderate to stronger relationships between course scores and End of Course tests.

This could be explained by the nature of the courses. The content for the U.S. History and Biology courses is more factual in nature and less skill based, requiring students to recall information. The Geometry course is skill based, however, students taking this course have typically been successful in the Algebra course. The grades for this course are not nearly as different from state averages as the scores for the Algebra course. The school's mean averages were 4 points behind the mean state averages for Biology, 14 points behind the mean state average for U.S. History, and 6 points behind the mean state average for Geometry. School scores are often compared to the state and district scores. In this school case it is the only school in the district; therefore the school often compares its score to schools with similar demographics.

Algebra had the weakest statistical relationship between end of course grades and EOCT scores. This was due to the significant difference between the Algebra course scores and those scores students earned on the Algebra End of Course Test. Course grade inflation may be explained by the fact that 2012-2013 was the first year of the revised curriculum as well as the

first administration of the new Common Core Georgia Performance Standards Algebra End of Course Test. The school experienced a 12% meets or exceeds pass rate on the End of Course Test, with a mean average of 61%. The state experienced a 41% meets or exceeds pass rate for the tests with the mean average of 67%.

The other two subjects, Ninth Grade Literature and Physical Science, had only weak to moderately positive relationships between end of course grades and EOCT scores . Typically the Ninth Grade Literature End of Course Test has the best overall scores of all tests examined for this study in 2011, 2010 and 2009 (GAOSA, Report Cards). The mean state average was 85%, while the study school's mean average was 82%. The Physical Science End of course Test mean average of 70% for the school was much lower than the mean state average of 89%. Both of these courses are skill-based courses, which may explain the strength of the relationships between course grades and EOCT scores for these classes.

The nature of the virtual school environment may also contribute to the differences between in the school studied and state averages. Barbour and Reeves (2009) discuss the attraction of virtual schools for students and parents due to the flexibility they offer. This flexibility combined with the non-traditional environment may also contribute to the differences between the virtual school's scores and states scores.

Grade inflation may also contribute to the differences in the scores. The number of As awarded to students has increased 28% since the 1960 at the college level (Rojstaczer & Healy, 2012). However, number of students needing remediation courses in reading and mathematics has also increased (Porter & Polikoff, 2011). Carifio and Carey (2010) attribute this to the policies put in place to help motivation and self-efficacy. Some of these policies include no zero policies and minimum grading polices. These policies are put into place because research shows

that students with low grades are not motivated but instead withdraw from learning (Guskey, 2011).

The second research question for this study was: is there a relationship between end of course grades and EOCT scores for students with disabilities? Separate Pearson's product-moment correlation analyses were run for each set, regular education students and those with disabilities. The results indicate a statistically significant relationship between end of course grades and EOCT scores for Algebra, Biology and U.S. History.

However, the results from the Ninth Grade Literature, Geometry, and Physical Science tests did indicate statistically significant results. The sample sizes for these analyses were small, ranging from 14 to 112 sets of data. Grade inflation may be the root cause for the lack of relationship between end of course grades and EOCT scores. End of Course pass rates are typically lower for students with disabilities at both at the school and state level compared to their regular education counterparts. Mastropieri *et al* (2006) reported similar findings in Virginia: 72% of students with disabilities passed the high-stakes test studied compared to 86% of Regular Education students.

Guskey and Jung (2009) express a concern over the need for meaningful grades for student with disabilities in regard to standards based classrooms. Many times students with disabilities are held to a completely different standard than their regular education peers. More emphasis is put on effort and completion than on quality and achievement. This causes grades for students with disabilities to be inflated. Students with disabilities are then at a disadvantage when forced to sit for the same standardized tests as their general education peers. However, as a result of NCLB and IDEA students with disabilities achievement scores have increased due to

increased participation in these tests (Schulte, Villwock, Whichard, & Stallings, 2001; Katsiyannis *et al*., 2007).

Numerous studies have been conducted in an attempt to find ways to assist students with disabilities pass high-stakes exams. Mastropieri *et al*. (2006) found that differentiating science instruction for students with disabilities increased student scores on both low-stakes (teacher made) and high-stakes standardized tests. The researchers conducted a the study over a 12 week period in which one group received the treatment/intervention of differentiation and one group received regular instruction. The study found that differentiated instruction in eighth-grade science classes scores was significantly higher for the experimental group, which received the intervention, on both unit and standardized high-stakes tests.

Carter, Wehby, Hughes, Johnson, Plank, Barton-Arwood, and Lunsford (2006) found an increase in scores for those students with disabilities who were taught test taking strategies. Carter *et al*. (2006) conducted a study on students with high-incidence disabilities in which the students were divided into two groups: One group received the high-stakes test strategies and the other group did not. Students were given a practice Tennessee Competency Achievement Program test, and scores of the two groups were compared with a paired-sample t test. While the implementation of test-taking strategies did help disabled students increase their scores, the increase did not equate to in all cases to achieving a passing score. Carter *et al*. (2006) suggests that the strategies be taught over a longer period of time and the implementation of strategies taught in earlier grades.

Certain instruction methods also increase the scores of students with disabilities. Steele (2010) suggests numerous strategies for assisting learning-disabled high school students prepare for high-stakes mathematics tests. These strategies included using real life example during

instruction, organizing instruction around themes, and providing organizational/ note taking tips. Scruggs, Mastropieri, and Okolo (2006) discuss many of these same strategies for the instruction of science and social studies. While these strategies are helpful in preparing students for high-stakes tests, accommodations level the playing field for students with disabilities.

The third research question for this study was: Is there an association between pass rates (earning a 70% or higher) in the course and pass rates (earning a 70% or higher) on the End of Course tests? Results from the McNemar chi squared test on paired proportion showed that there is a statistically significant association between passing a course and passing the End of Course Test for all subjects tested except Ninth Grade Literature. There were 91 students who failed the Ninth grade literature course but passed the Ninth Grade Literature End of Course test and 107 students who passed the course and failed the test. Most of the other courses studied had very few students who passed the End of Course Test and failed the course. This could be explained by the nature of the course, by the consistency in grading by the teachers, or by the potentially elementary nature of the test itself. More investigation is needed to explain why Ninth Grade Literature end of course grades and EOCT scores are not more closely associated. These results suggest that the issue is passing courses rather than relationship between high states tests and graduation, as Jonsson (2001), Nichols, Glass, and Berliner (2006), and others suggest.

The virtual school underperformed by state standards on four of the six courses studied. The school outperformed the state in Ninth Grade Literature and Biology. In Algebra, 89% of students earned a passing score in the course and only 12% of the students passed the End of Course Test associated with that course. This was an 18% difference from the percent the State of Georgia reported as passing the End of Course Test in Algebra in 2013. In Geometry, 83% of students passed the course, but only 57% of the students passed the End of Course Test

associated with that course. This was a difference of 19% from the percent the State of Georgia reported passing Geometry in 2013. In Physical Science, 72% of students passed the course, but only 51% passed the End of Course Test associated with that course. This was a difference of 32% from the percent that the State of Georgia reported passing Physical Science in 2013. In U.S. History, 93% passed the course but only 46% of students passed the End of Course Test in US History. This was a 27% difference from the state of Georgia reported 73% passing in 2013. Ninth Grade Literature and Biology were the only two courses to where the school studied out preformed the state of Georgia reported average (EOCT State Scores).

The nature of the schools at risk population may explain the difference between the state passing percentages and the schools passing percentages. The 1983 *A Nation at Risk* defined students who are in danger of academic failure as at-risk due to low socioeconomic status, below average grades, or in danger of grade retention (Placier, 1993; Spring, 2010). Jacob (2001) reports that students in the lower 25th percentile are more likely to drop out of school than their counterparts in higher percentiles because they feel pressured by high-stakes tests. Paris and McEnvoy (2000) attribute poor test scores to student motivation. Students who do not feel that the tests are valid will not be motivated to take the tests seriously or do well. Amrien and Berliner (2003) indicate that there was no direct correlation between increased motivation and standardized tests.

Limitations

Four limitations of this study have been determined and will be discussed below. Understanding the limitations of a study is important when considering its validity.

The first limitation of the study is the setting from which the data were collected. The gathered data came from a virtual high school that serves approximately 2,500 students around

the State of Georgia. These students attend school in a virtual environment. State certified teachers align a national curriculum to match the state identified standards. The school embraces a mastery based approach, meaning students earn credit for a course by demonstrating mastery of the curriculum. Students do not meet with the teacher for the courses on a daily basis. In fact, attending live class sessions is not a requirement at all. Students can self-pace but are given a suggested calendar of assignments for each semester. Students are encouraged to finish the course work for a course prior to sitting for the End of Course Test.

The second limitation of the study is the large sample sizes for each test. The sample size of the tests ranged from 419 for U.S. History to 968 in Ninth Grade Literature. As the sample sizes increase, so does the likelihood of statistical significance increase (Rummel, 1976). So while the results of the Pearson product-moment correlation analyses indicate results of statistical significance, the effect sizes for most analyses range from weak to moderate. Only one of the whole group analyses indicate effect sizes of moderate to strong, U.S. History (Cohen, 1988). Effect size is defined by Gall *et al*. (2007) as the statistical measure of strength of an observed distance between two groups on a test or other instrument or the strength of an observed relationship between two or more measured variables. The practical significance of these tests due to the weak to moderate effect size indicates uniqueness to the data analyzed and is not indicative to all standardized tests.

The third limitation is the size of the sample for the Pearson product-moment correlation tests conducted on the data from students with disabilities. Sample sizes from three of the six courses studied had small sample sizes. The sample sizes for students with disabilities were 14 for Biology, 29 for U.S. History, and 38 for Geometry. Gall *et al*. (2007) recommend a minimum of 30 participants for correlational research. The analyses of two of the three groups,

Biology and U.S. History, indicate significant positive correlations with moderate to strong effect sizes.

The fourth limitation of the study is the non-tenable assumptions of linearity and homoscedasiticity of the scatterplots. The scatterplots are noisy with multiple outliers and therefore do not look like the textbook examples of scatterplots (Gall *et al*., 2007). As a result, the data should be interpreted cautiously. The correlations may slightly underestimate or overestimate the strengths of the correlations. This may influence the practical significance of the data for making future policy.

Implications

Despite the limitations, there is much that can be learned from the findings of this study. While not all students at this school are finding success on state-mandated standardized tests, this research affirms that there is some relationship between success in the class and success on the state mandated End of Course Test. Typically students who earn higher grades in the course did in turn earn higher grades on the End of Course Test for that course. However, there are still gains that need to be made in order for the school to meet the CCRPI goals for achievement on standardized tests and to meet the state of Georgia reported passing percentages for each End of Course Test. CCRPI evaluates schools on the number of students based on mastery, which is defined by the State of Georgia as meeting or exceeding standards on the End of Course tests associated with Ninth Grade Literature, American Literature, Algebra, Geometry, Biology, Physical Science, U.S. History, and Economics (Barge, 2013).

The subjectivity of grading practices of teachers could explain the results from these analyses. More students are passing the courses than pass the End of Course Test for that subject even though the analyses indicate there is a positive relationship between passing the course and

passing the test associated with that course. The virtual school and other schools need to come up with strategies for increasing the percentage of students passing the tests. Relying more on summative assessments, rather than formative assessments may close the gap between passing the course and failing to meet standards on the End of Course tests. There is also some room for growth and conversation in the areas of Ninth Grade Literature, where many students earned failing end of course grades but passed the End of Course Test, and in Algebra where many students passed the course with high scores but failed the End of Course Test.

Recommendations

Additional research needs to be conducted to further understand the relationship between course scores given by teachers and grades on standardized tests associated with those courses. Suggestions for additional studies related to this study include but are not limited to the following:

1. Expanding this study to include all virtual schools.
2. Expanding the study to include all high schools in the state that have similar demographics as the school studied.
3. Replicating this study in other states that use standardized tests to assess student knowledge of a course as part of the overall course grade.
4. Expanding the study to include high schools with demographics that differ from the school studied.

REFERENCES

Abrams, L. (2004). Teachers view on high-stakes testing: Implications for the classroom policy brief [electronic version]. *Education Policy Studies Laboratory*. Arizona State University, 1-36.

Achieve Inc.: The American Diploma Project (2004). *Do graduation tests measure up?: A closer look at state high school exit exams.* retrieved on March 8, 2007, from http://achieve.org/node/581.

Amrein, A. & Berliner, D. (2003). The effects of high-stakes testing on student motivation and learning [electronic version]. *Educational Leadership 60*(5), 32-38.

Anyon, J. & Greene, K. (2007). No child left behind as an anti-poverty measure. *Teacher Education Quarterly 34*(2), 157-162.

Archambault, L. & Kennedy, K. (2012). Offering preservice teachers field experience in k-12 online learning: A national survey of teacher education programs. *Journal of Teacher Education 62* (3), 185-212.

Au, W. (2009). Social studies, social justices: W(h)ither the social studies in high-stakes testing? *Teacher Education Quarterly 36*(1), 43-58.

Barbour, M. & Reeves, T. (2009). The reality of virtual schools: A review of the literature. *Computer & Education 52*(2), 402-416.

Barge, J. (2013). *College & Career Ready Performance Index*. Georgia Department of Education, CCRPI Division. Accessed 1 January 2014 http://www.gadoe.org/CCRPI/Pages/default.aspx

Barge, J. D. (2014). Georgia's Race to the Top (RT3) Plan. Georgia Department of Education. Retrieved from http://www.gadoe.org/Race-to-the-Top/Pages/default.aspx.

Berliner, D. C. (2009). MCLB (much curriculum left behind): A U.S. calamity in the making. *The Educational Forum 73*(4), 284-296.

Borman, G. D. & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational 78*(3), 367-409.

Bracey, G. W. (2007). The first time 'everything changed.' *Phi Delta Kappan 89*(2), 119-136.

Brady, K., Umpstead, R., Eckes, S. (2010). Unchartered territory: The current legal landscape of public cyber charter schools. *B.Y.U. Education and Law Journal* 10(2), 191-248.

Braun, H. (2004). Reconsidering the impact of high-stakes testing, *Education Policy Analysis Archives, 12* (1) 1-74.

Byrnes, M. (2009). *Taking sides: Clashing views in special education* (4th ed.). McGraw-Hill.

Carifio, J. & Carey, T. (2009). A critical examination of current minimum grading policy recommendations. *The High School Journal 3(*1), 23-37.

Carifio, J. & Carey, T. (2010). Do minimum grading practices lower academic standards and produce social promotions? *Education Horizons 88*(4), 219-230.

Carter, E. W., Wehby, J., Hughes, C., Johnson, S. M., Plank, D. R., Barton-Arwood, S., & Lunsford, L. B. (2006). Preparing adolescents with high-incidence disabilities with strategy instruction. *Preventing School Failure 49*(2), 55-62.

Christenson, S., Decker, D., Trizenberg, H., Ysseldyke, J. & Reschly, A. (2007) Consequences of high-stakes assessment for students with and without disabilities. *Educational Policy 21* (4) 662-690.

Clarke, M., Haney, W., and Madaus, G. (2000). High-stakes Tests and High School

Completion [electronic version] *NBETTP Statements 1*(3), 1-12.

Crocco, M. S. & Costigan, A. T. (2007). The narrowing of curriculum and pedagogy in the age of accountability. *Urban Education 42*(96), 512-535.

Cohen, J. (1988) Statistical Power Analysis for Behavioral Sciences. Lawerence Erlbaum Associates: New Jersey, 590.

Cox, K. (2006) What Georgia educators need to know about Georgia's testing program. (Brochure). *Georgia Department of Education, Testing Division*. Accessed 15 July 2010 from http://dlgmaint.galib.uga.edu/cgi-bin/govdimag.cgi?user=public&page=1&path=2006/ga/e300_pt45/m1/2006/t4/folio.con&return=http%3A%2F%2Fdlgmaint.galib.uga.edu%2Fcgi-bin%2Fgovdimag.cgi%3Fpage%3D6%26path%3D2006%2Fga%2Fe300_pt45%2Fm1%2F2006%2Ft4%2Ffolio.con

Cox, K. B. (2011). Putting classroom grading on the table: A reform in progress. *American Secondary Education 4*(1), 67-87.

Dee, T. (2003). Learning to earn. *Education Next 3*(3), 1-3.

Distance education in public high schools (indicator 15-2012) *The National Center of Educational Statistics.* Accessed on 4 April 2013 from http://nces.ed.gov/programs/coe/indicator_ths.asp

Ellis, C. (2007). No child left behind: A critical analysis. *Curriculum and Teaching Dialogue 9*(1-2), 221-233.

EOCT State Scores. (2014). Georgia Department of Education Website. Retrieved 1 March 2014 from http://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Pages/EOCT-Statewide-Scores.aspx

Finnigan, K. & Gross, B. (2007) Do accountability sanctions influence teacher motivation? Lessons from Chicago's low-performing schools. *American Education Research Journal 44*(3), 594-629.

Forum guide to elementary/secondary virtual education. *National Center for Education Statistics*. Accessed on 4 April 2013 from www.nces.ed.gov/pubs2006/2006803.pdf

Gall, M., Gall, J., & Borg, W. (2007) *Education Research: An Introduction*. Boston: Pearson, 672.

GA State Board Rule 160-4-2-.13 (2006) Accessed 07 March 2014 from http://archives.gadoe.org/pea_board.aspx?PageReq=PEABoardRules

Germain, M. & Scandura, T. (2005) Grade inflation and student individual differences as systematic bias in faculty evaluations. *Journal of Instructional Psychology 32*(1), 58-67.

Gerwin, D. & Visone, F. (2006) The freedom to teach: Contrasting history teaching in elective and state--tested courses. *Theory and Research in Social Education 34*(2), 259-282.

Godfrey, K. (2011). Investigating grade inflation and non-equivalence. Research report 2011-2. College Board. Accessed 23 March 2013 from http://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-2011-2-investigating-grade-inflation-non-equivalence.pdf

Goodwin, B (2011). Grade inflation: Killing with kindness? *Educational Leadership 69*(3), 80-81.

Gordan, M. & Fay, C. (2010) The effects of grading and teaching practices on students' perceptions of grading fairness. *College Teaching 58*(3), 93-98.

Governor's Office of Student Achievement. State of Georgia. Online Reports, Report Card

retrieved 5 October 2013 from www.gaosa.org.

Greene, J. & Forster, G. (2003). Public high school graduation and college readiness rates in the United States. *Manhattan Institute for Policy Research.* Accessed on 7 April 2013 from http://www.manhattan-institute.org/html/ewp_03.htm

Greene, J. & Winters, M. (2004). Pushed out or pulled up? Exit exams and dropout rates in public high schools, *Education Working Paper: Manhattan Institute for Policy Research 5*. Accessed on 13 March 2013 from www.manhattah-institute.org/html/ewp_05.htm

Griffin, B. & Heidorn, M. (1996). An examination of the relationship between minimum competency test performance and dropping out of high school *Educational Evaluation and Policy Analysis 18,* 243-252.

Guskey, T. (2006). Making high school grades meaningful. *Phi Delta Kappan 87*(9), 670-675.

Guskey, T. (2011). Five obstacles to grading reform. *Educational Leadership 69*(3), 16-21.

Guskey, T (2011). Stability and change in high school grades. *National Association of Secondary School Principles 95*(2), 85-98.

Guskey, T. R., and L. A. Jung. 2009. Grading and reporting in a standards-based environ-ment: Implications for students with special needs. *Theory Into Practice*48(1): 53–62.

Haney, W., Madaus, G., Abrams, L., Wheelock, A., Miao, J., & Gruia, I. (2004). The education pipeline in the United States 1970-2000 [electronic version]. *NBETPP*.

Howell, J. (2011). What influences students need for remediation in college? Evidence from California. *Journal of Higher Education 82*(3), 1-28.

Hunt, J. *et al*. (2009). The impact of the pressures to make adequate yearly progress on teachers in a Midwest urban school district: A qualitative analysis. *Journal of Thought 44*(3/4), 63-75.

Jacob, B. (2001). Getting tough? The impact of high school graduation exams. *Educational Evaluation and Policy Analysis 23*(2), 99-121.

Johnston, S. & Barbour, M. (2013). Measuring success: Examining achievement and perceptions of online advanced placement students. *The American Journal of Distance Education 27*(1), 16-28.

Jonsson, P. (2001). Higher standards-and more dropouts?: More North Carolina students are quitting school, and some say new 'high-stakes' tests may be a factor [electronic version]. *Christian Science Monitor 93*(119), 3.

Journell, W. (2010). The influence of high-stakes testing on high school teachers' willingness to incorporate current political events into the curriculum. *The High School Journal 93*(3), 111-125.

Kamber, R. & Biggs, M. (2003). Grade inflation: Metaphor and reality. *Journal of Education 184*(1), 31-37.

Katsiyannus, A., Zhang, D., Ryan, J. B. & Jones, J. (2007) High-stakes testing and students with disabilities: Challenges and promises. *Journal of Disabilities Studies 18* (2) 160-167.

Kezim, B., Pariseau, S., & Quinn, F. (2005) Is grade inflation related to faculty status? *Journal of Education for Business 80*(6), 358-363. .

Madaus, G. & Clarke, M. (2001). The adverse impact of high-stakes testing on minority students: Evidence from 100 years of test data, High Stakes K - 12 Testing Conference, Harvard University.

Martone, A., & Sireci, S. (2010). Evaluating alignment between curriculum, assessment and instruction. *Review of Educational Research 79*(4), 1332-1361.

Mastropieri, M., *et al*. (2005). Case studies in co-teaching in the content areas: Successes, failures and challenges. *Intervention in School and Clinic 40*(5), 260-270.

Mastropieri, M. A., Scruggs, T. E., Norland, J.J., Berkeley, S., McDuffie, K., Tornquist, E. H. & Conners, N. (2006). Differentiated curriculum enhancement in inclusive middle school science: Effects on classroom and high-stakes tests. *The Journal of Special Education 40(*3), 130-137.

Maxcy, B. D. (2011). The politics of priorities in turbulent times: Policy logics, faces of power, and reform possibilities. *Peabody Journal of Education 86*(3), 252-271.

McFarlane, D. (2011). Are there differences in the organizational structure and pedagogical approach of virtual and brick-and-mortar schools? *Journal of Multidisciplinary Research 3*(2), 83-98.

McGuinn, P. J. (2006) No *child left behind and the transformation of federal policy, 1965-2005*. Lawrence, Kansas: University of Kansas Press, 260.

McNeil, L. M., Coppola, E., Radigan, J., & Vasquez, H. (2008). Avoidable losses: High-stakes accountability and the dropout crisis. *Education Policy Analysis Archives, 16 (3).* Retrieved 12 September 2010 from http://epaa.asu.edu/epaa/v16n3/.

Nichols, S., Glass, G., & Berliner, D. (2006). Much pain, but no gains [electronic version]. *NEA Today*. *24*(4), 10.

Nichols, S. L. & Berliner, D. C. (2008) Testing the joy out of learning. *Educational Leadership 65*(6), 14-18.

Nowicki, S., Duke, M., Sisney, S., Sticker, B., & Tyler, M.A. (2004). Reducing the drop-out rate

of at-risk high school students: The effective learning program (ELP). *Genetic, Social & General Psychology Monographs 130*(3), 225-239.

Pace, J. (2008) Inequalities in history-social science teaching under high-stakes accountability: Interviews with fifth-grade teachers in California. Social Studies Research and Practice 3 (1) 24-40.

Paris, S. & McEnvoy, A. (2000). Harmful and enduring effects of high-stakes testing. *Issues in Education. 6*(1/2), 145.

Paris, S. & Urdan, T. (2000). Polices and practices of high-stakes testing that influence teachers and schools. *Issues in Education 6*(1/2), 83.

Pedulla, J., Abrams, L., Madaus, G., Russel, M., Ramos, M., & Miao, J. (2003). Perceived effects of state mandated testing programs on teaching and learning: Findings from a national survey of teachers [electronic version]. Chestnut Hill, MA: NBETPP, Boston College.

Picciano, A. & Seaman, J. (2007). K-12 online learning: A survey of US school district administrators. *Babson Survey Research Group*. Accessed on 4 April 2013 from http://sloanconsortium.org/publications/survey/K-12_06

Placier, M.L. (1993). The Semantics of Policy Making: The Case of "At Risk." *Educational Evaluation and Policy Analysis*, *15*(2), 380.

Porter, A. & Polikoff, M. (2011). Measuring academic readiness for college. *Educational Policy 26*(3), 394-417.

Randall, J. & Englehard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education 26*(7), 1372-1380.

Reid, K., Aqui, Y., & Putney, L. (2009). Evaluation of an evolving virtual high school.

Educational Media International 46(4), 281-294.

Roblyer, M. (2006). Virtually successful: Defeating the dropout problem through online school programs. *Phi Delta Kappan 88*(1), 31-36.

Rojstaczer, S. & Healy, C. (2012). Where A is ordinary: The evolution of American college and university grading, 1940-2009. *Teachers College Record 114*(7), 1-23.

Rose, R. & Blomeyer, R. (2007). Access and equity in online classes and virtual schools. *North American Council for Online Learning*. Accessed 4 April 2013 from http://www.k12onlineresearch.org/index.php/P10

Rummell, R. J. (1976) Understanding Correlation. Honolulu: University of Hawaii Press. [electronic version]. http://www.hawaii.edu/powerkills/UC.HTM

Ryan, R. & Weinstein, N. (2009). Undermining quality teaching and learning: A self-determination theory perspective on high-stakes testing. *Theory and Research in Education 7*(2), 224-233.

Schulte, A. C., Villwock, D. N., Whichard, S. M., & Stallings, C. F. (2001) High-stakes testing and expected progress standards for students with learning disabilities: A five-year study of one district. *School Psychology Review 30* (4) 487-506.

Scuggs, T., Mastroperi, M., Okolo, C. (2006). Science and social studies for students with disabilities. *Focus on Exceptional Children 41*(2), 1-21.

Slavin, R. (1997). *Education Psychology: Theory and Practice*. Allyn and Bacon: Boston, 612.

Spitler, C., Repetto, J. & Cavannaugh, C. (2013). Investigation of a special education program in a public cyber charter school. *The American Journal of Distance Education 27*(1), 4-15.

Spring, J. (2010). *American Education*. Boston: McGraw-Hill, 320.

Steele, M. (2010) High school students with learning disabilities: Mathematics instruction, study skills, and high-stakes tests. American Secondary Education *38*(3), 21-27.

Stillman, J. (2009) Taking back the standards: Equity-minded teachers' responses to accountability related instruction restraints. The New Educator 5 (2) 135-160.

Sunderman, G. (2009). The federal role in education: From the Reagan to the Obama administration. *Voices Urban Education 24,* 1-9.

Syverson, M. (2009). Social justice and evidence-based assessment with the learning record. *Forum on public policy online*. Accessed 24 September 2010 from http://eric.ed.gov/ERICWebPortal/search/simpleSearch.jsp?newSearch=true&eric_sortField=&searchtype=keyword&pageSize=10&ERICExtSearch_SearchValue_0=EJ864839&eric_displayStartCount=1&_pageLabel=ERICSearchResult&ERICExtSearch_SearchType_0=kw

U.S. Department of Education. (2009b). *Race to the Top program: Executive summary*. Washington, DC: Accessed 24 September 2010 from http://www2.ed.gov/programs/racetothetop/executive-summary.pdf

Van Brummelen, H. (2002). *Steppingstones to curriculum*. Colorado Springs: Purposeful Design Publications.

Walden, L. and Kritsonis, W. (2008). The impact of the correlation between the No Child Left Behind Act's high-stakes testing and the high drop-out rates of minority students. *National Journal for Publishing and Mentoring Doctoral Student Research 5*(1), 1-6.

Warren, John Robert & Grodsky, Eric. (2009). Exit exams harm students who fail them—and don't benefit students who pass them. *Phi Delta Kappan 90*(9), 645-649.

Watson, J., Murin, A., Vashaw, L., Gemin, B. & Rapp, C. (2012) Keeping pace with k-12 online & blended learning: An annual review of policy and practice. *Evergreen Education Group.* Accessed 4 April 2013 from http://kpk12.com/reports.

Wills, J. S. (2007). Putting the squeeze on social studies: Managing teacher dilemmas in subject areas excluded from state testing. *Teachers College Record 109*(8), 1980-2046.

Wills, J. S. & Sandholtz, J. H. (2009). Constrained professionalism: Dilemmas of teaching in the face of test-based accountability. *Teachers College Record 111*(4), 1065

Wineburg, S. (2004). Crazy for history. *The Journal of American History 90*(4), 1401-15.

Wongsurawat, W. (2009). Does grade inflation affect the credibility of grades? Evidence from US law school admissions. *Education Economics 17*(4), 523-534.

Woodruff, D. & Ziomek, R. (2004). High school grade inflation from 1991-2003. *ACT Research Report Series.* Accessed 19 March 2013 from www.act.org/research/researchers/reports/pdf/ACT_RR2004-4.pdf

Yaffe, D., Coley, R., & Pliskin, R., eds. (2009). Addressing achievement gaps, national policy and innovations. *ETS Policy Notes 17*(1), 1.

Ysseldyke, J., Nelson, J. R., Christenson, S., Johnson, D. D., Dennison, A. Triezenberg, H., Sharpe, M. & Hawes, M. (2004) What we know and need to know about the consequences of high-stakes test for students with disabilities. *Council for Exceptional Students 71* (1) 75-91.

Ziomek, R. & Svec, J. (1997). High school grades and achievement: Evidence of grade inflation. *NASSP Bulletin 81*(587), 105-108.