POLICY IMPLICATIONS OF SELECT STUDENT CHARACTERISTICS AND THEIR INFLUENCE ON THE FLORIDA BIOLOGY END-OF-COURSE ASSESSMENT

By

JANINE CECELIA BERTOLOTTI

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF EDUCATION

UNIVERSITY OF FLORIDA

2013

To my parents,
Michael and Orlene Ireland,
my husband,
Joshua Bertolotti,
and my daughter,
Brooke Cecelia Bertolotti

ACKNOWLEDGMENTS

"I can do all things through Christ who strengthens me"
- Philippians 4:13

This endeavor would be unsuccessful had it not been for the encouragement of family, friends, and colleagues. I would like to thank my parents, Michael and Orlene Ireland, for encouraging me in my academic pursuits and the endowment of financial support at both the undergraduate and graduate level. My brother, Garth Ireland, who has always been my cheerleader and friend. Most of all, my success has been dependent on the unwavering love and support of my beloved husband, Joshua Bertolotti. Your support has been invaluable, despite the subsequent imposed move and deployment that accompanied it. I am so thankful we have made it through this journey together and am very grateful that you will always be in my life. Last, I thank my daughter, Brooke Cecelia Bertolotti, whose recent addition within the past year has broadened my perspective on life and has served as a motivator to complete this venture. I am forever indebted to each of you, especially my husband.

Dr. Bernard Oliver, my doctoral committee chair and advisor, has provided helpful guidance and patience throughout this process, ensuring that this study would be completed in time for the execution of military orders. Drs. Wood and Eldridge have provided direction to my cohort throughout the past three years, challenging us and guiding us throughout this entire process. Dr. Adams has also been valuable in sharing of both time, resources, and expertise on my committee. Dr. Mousa for always supporting me and providing me with feedback during my entire academic life at Florida. Collectively, you have all played a role in shaping my personal and academic growth as a researcher and educator at The University of Florida. Thank you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACHIEVEMENT LEVEL	This defines the level of success a student has with the objectives of the curriculum dictated by the *Next Generation Sunshine State Standards* (NGSSS). They range from level 1 (inadequate level of success) to level 5 (demonstration of mastery). Level 3 is considered a satisfactory level of success. Achievement levels exist for FCAT 2.0 Reading but not for Biology I end-of-course assessments which will be developed in Spring 2013.
END-OF-COURSE	A comprehensive test taken during the academic year that covers the objectives of the curriculum derived from the *Next Generation Sunshine State Standards* (NGSSS).
FCAT 2.0	The newly revised version of Florida Comprehensive Assessment Test (FCAT) that measures student achievement based on content standards derived from the *Next Generation Sunshine State Standards* (NGSSS).
TITLE 1 SCHOOL	This refers to a school receiving Federal Title 1 funds which provides financial assistance to schools with high numbers or high percentages of children from low-income families to help ensure that all children meet challenging state academic standards.

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Education

POLICY IMPLICATIONS OF SELECT STUDENT CHARACTERISTICS AND THEIR INFLUENCE ON THE FLORIDA BIOLOGY END-OF-COURSE ASSESSMENT

By

Janine Cecelia Bertolotti

December 2013

Chair: Bernard Oliver
Major: Educational Leadership

In an attempt to improve student achievement in science in Florida, the Florida Department of Education implemented end-of-course (EOC) assessments in biology during the 2011-2012 academic school year. Although this first administration would only account for 30% of the student's overall final course grade in biology, subsequent administrations would be accompanied by increasing stakes for students, teachers, and schools. Therefore, this study sought to address gaps in empirical evidence as well as discuss how educational policy will potentially impact on teacher evaluation and professional development, student retention and graduation rates, and school accountability indicators.

This study explored four variables- reading proficiency, ethnicity, socioeconomic status, and gender- to determine their influence and relationship on biology achievement on the Biology I EOC assessment at a Title 1 school. To do so, the results of the Biology I EOC assessment administered during the Spring 2012 school year was obtained from a small, rural Title 1 high school in North Florida. Additional data regarding each student's qualification for free and reduced-price lunch, FCAT Reading developmental scale scores, FCAT Reading level, grade level, gender, and ethnicity

were also collected for the causal-comparative exploratory study. Of the 178 students represented, 48% qualified for free and reduced-price lunch, 54% were female, and 55% scored at FCAT Reading level 3 or higher. Additionally, 59% were White and 37% Black.

A combination of descriptive statistics and other statistical procedures such as independent samples one-tailed *t*-test, one-way ANOVAs, ANCOVAs, multiple-regression, and a Pearson *r* correlation was utilized in the analysis, with a significance level set at 0.05. Results indicate that of all four variables, FCAT Reading proficiency was the sole variable, after adjusting for other variables; that had a significant impact on biology achievement. Students with higher FCAT Reading developmental scores scored significantly higher on the Biology I EOC assessment than their peers with lower FCAT Reading scores. Additionally, FCAT Reading developmental scale scores were significantly correlated with Biology I EOC scores. The significant predictors for biology scores included FCAT Reading developmental scale scores, grade level, and eligibility for free lunch, which collectively explained 60% of the variability.

CHAPTER 1
INTRODUCTION

Historical Context of School Accountability

Since the 1983 publication of *A Nation at Risk* urged greater accountability from schools and higher academic standards (National Commission on Excellence in Education, 1983), the standards-based reform movement has gained momentum and exerted considerable influence on both state and federal educational policy. Dismal student achievement data on national and international assessments served as a catalyst for the exigency of school reform. One such reform, the No Child Left Behind Act (NCLB) was signed into law on January 8, 2002, by President George W. Bush. NCLB was based on the premise that by holding America's public schools accountable for student performance, student achievement would increase. This reform movement assumed that the foundation for academic achievement rests primarily within the public schools whose failure to educate American's children in basic reading and math literacy resulted in a future citizenry that was unprepared for employment and advanced educational opportunities.

The proposed solution involved ensuring school accountability through: educational standards in each grade and subject, testing, and accountability through standardized assessments in core subject areas. Despite statewide variations in policy, it was common for states to create content standards in the core academic subject areas and test regularly to measure mastery of standards. These tests were developed by each state and approved at the federal level for public schools to obtain federal funding. Some states awarded special diplomas for students who scored exceptionally

high, while students receiving low scores might be retained or denied a diploma (American Educational Research Association, 2000).

The standardized tests were also used to supply objective performance indicators that determined whether schools made "adequate yearly progress" in educating their students. The scrutiny of published test results served to ensure that administrators and teachers would improve educational delivery in a manner that would prevent "failing" and achieve the rewards associated with high achievement (Carnoy & Loeb, 2002). High scores could bring public praise or financial rewards; low scores could bring public embarrassment and heavy sanctions. Failure to make adequate progress over a period of time resulted in the implementation of sanctions such as allowing students the option to matriculate at other better performing private or public schools (No Child Left Behind Act of 2001, 2002).

Prior to the prominence of the accountability movement, the association between assessment and accountability was laissez-faire (Carnoy & Loeb, 2002; Kress, Zechmann, & Schmitten, 2011). Accountability was traditionally community oriented and relied on the participation of parents and local school boards. Schools were "accountable to district administrators, who, in turn, answer to elected boards" (Carnoy & Loeb, 2002, p.306). Assessments were primarily used by administrators and teachers for diagnostic purposes, to determine student proficiency with loosely defined state curricula, and to segregate students into academic tracks (Carnoy & Loeb, 2002; Kress, Zechmann, & Schmitten, 2011). However, after NCLB, the competency of teachers was no longer the primary judge of student accomplishment; instead, it became increasingly commonplace to rely on testing (Baker, 2001).

This routine use of tests has gained traction due to escalating public views of test credibility as the primary source of information regarding individual student performance, the quality of educational institutions, and the quality of innovations in education (Baker, 2001). Marchant (2004 identified a shift from testing being used for diagnostic/ prescriptive purposes that allow teachers to evaluate student proficiency to becoming a yardstick to evaluate the success of students, teachers, schools, districts, and even states. Thus testing has become a high-stakes process in which achievement tests are associated with serious consequences for educators and schools.

Education Policy

Public educational policy is based on the premise "that all children... obtain a high-quality education and reach a minimum proficiency on challenging state academic achievement standards and state academic assessments" (U.S Congress, 2002, p. 15). In Florida, this public policy has been adopted through the implementation of the Next Generation Sunshine State Standards (NGSSS) and an increase in the minimum passing requirements for graduation, as well as the adoption of Common Core standards. Common Core standards are English and mathematics educational standards for kindergarten through 12th grade that had been adopted by 45 states, including Florida by the end of 2010 (Florida Department of Education, 2010) (Common Core Standards Initiative, 2012a). However, these standards will not be fully implemented in Florida until the 2013-2014 school year (Florida Department of Education, 2010).

The Common Core standards are considered to promote equity by ensuring that all students, regardless of geographic location, are provided with the same clear set of expectations for knowledge and skills at each grade level (Common Core Standards

Initiative, 2012a). The standards are designed to be "robust and relevant to the real world, reflecting the knowledge and skills that our young people need for success in college and careers" so that students can effectively compete locally and globally (Common Core Standards Initiative, 2012c). According to the Common Core Standards Initiative (2012b), Common Core standards use: increasing levels of complexity in student reading and reading comprehension; writing standards heavily focused on developing and conveying logical arguments based on substantive claims and sound reasoning; techniques to develop student vocabulary; and provide for a solid foundation in whole numbers, addition, subtraction, multiplication, division, fractions, and decimals. Thus, collectively the Common Core and the Next Generation Sunshine State Standards serve as additional new accountability indicators that strive to improve the academic achievement of all students. This focus demonstrates how Florida's education policy centers on improving the rigor, skills, and core content knowledge through the implementation of increasing number of standards and assessments to measure student learning.

Although the Common Core is meant to improve student performance through systematic incorporation of core standards in reading and math, the extent of the effect of this policy on science education is subject to debate. Education policies such as Common Core and NCLB, which do not specifically target science education, still exert an influence. Johnson (2013) revealed that the significant emphasis placed on reading and math had resulted in significant reductions in the emphasis on science education. For example, middle schools were burdened with trying to bridge the deficits in student knowledge and experiences due to the failure of elementary schools to teach science.

Therefore, the practices within science education are influenced by education policy, especially ones specifically targeted for science education (Fensham, 2009). This influence is in part, due to the endorsement of some practices over others as well as the differential emphasis of the *what* and *how* of science teaching and learning. Thus, education policy, "are operational statements of values" (Fensham, 2009, p. 1080).

One such example is Thailand's policy decision to increase student achievement in the sciences. The policy required high school students focusing on sciences, to study physics, chemistry, and biology for three years each. Even students who focused on the humanities, were required to complete 2 years of physical and biological sciences. This policy of "no choice" resulted in the reduction in gender inequity in student performance in the sciences (Fensham, 2009). Despite this positive outcome, such a policy decision within the United States may not be a national possibility. The strong cultural emphasis in American education on the provision of "choice" may inadvertently result in less choices at the postsecondary level and future careers (Fensham, 2009).

Thus the effect of policy on practice is a significant one. This effect is particularly pronounced in federally mandated school accountability situations that has influenced school funding and district strategic planning (Johnson, 2013). As state developed multiple-choice assessments have increased in popularity, emphasis on student gains on such assessments has become top priority, thereby affecting funding, personnel, scheduling, curriculum and instruction, the learning environment, and accountability (Johnson, 2013). The collective effect of both the macro- and micro- level education policy produces educational turbulence and stifles systemic educational reform

(Johnson, 2013), as long as test-based accountability systems maintain their current theoretical underpinnings (Anderson, 2011).

According to Anderson (2011), "accountability policies alter the practice of science education because they limit the effectiveness of other reform movements" (p. 116). The use of innovative practices such as project-based learning, student-centered teaching, inquiry-based instruction, and constructivist learning; that research has identified as improving student engagement and understanding of science; are typically sidelined to focus on improving standardized test scores. Research has also revealed that due to this accountability test system, the highest performing students may have received a reduction in attention and opportunities due to teacher focus on strategically improving the scores of a few students to pass state tests (Anderson, 2011). Furthermore, the availability and funding of professional development for teachers in science was also negatively impacted, so as to address the accountability requirements in math and reading (Anderson, 2011).

Therefore, although expectations of science education nationally and state-wide are high, there exists a disconnect in what takes place in America's classrooms. Thus, education policy as it relates to the curriculum and measurement of student learning, has served to alter science education; in part, due to its significant emphasis on student achievement as indicated by standardized assessments. Nevertheless, tensions exist within the public educational policy community about current approaches to the measurement of learning (Baker, 2001). Public policy is shaped by three competing desires: a) the creation of clear goals and appropriate measures so that students, teachers, and schools can meet such expectations, b) local control in deciding what the

goals and measures should be, and c) using the least expensive way of comparing students, schools, districts, and states (Baker, 2001). According to Baker (2001), these three conflicting goals lead to the reliance on inappropriate tests to measure progress, and thus public policy promulgates faulty and premature conclusions from the tests.

Factors Affecting Student Achievement Scores on Standardized Tests

No test is capable of perfectly measuring the knowledge of a student, hence why educators have consistently expressed concerns about the conclusions derived from standardized testing. In fact, test scores fail to reveal other factors that influence student achievement. Research has revealed a variety of other characteristics that have a negative impact on student performance on standardized tests include, but are not limited to: low levels of reading proficiency (Haught & Wall, 2004; Maerten-Rivera, Myers, Lee, & Penfield, 2010; O'Reilly & McNamara, 2007; Romance & Vitale, 2008), limited background knowledge (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Ozuru, Dempsey, & McNamara, 2009), limited vocabulary (Fang, 2006), negative attitudes toward science (Singh, Granville, & Dika, 2002), and student participation in less academically rigorous courses (Moore & Slate, 2008; National Science Foundation, 2012a; Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Rumberger & Palardy, 2005; Strand, 2012).

Research has also revealed that student backgrounds have a significant impact on student achievement scores. However, that impact is not a simple direct product of social background, but instead reveals itself indirectly through complex stratification mechanisms (Mostafa, 2010). For example, research has revealed that students from families with lower levels of parental educational attainment as well as single parent families, have lower levels of educational achievement (Hines& Holcomb-McCoy, 2013;

Kahlenberg, 2001; Okpala, Smith, Jones, & Ellis, 2000; Qui & Wu, 2011). Research also shows that these families are more likely to have lower earning potential (National Center for Children in Poverty, 2007; U.S. Department of Labor (2010). Since the socioeconomic composition of the local community largely determines the socioeconomic composition of schools, schools are usually composed of other individuals of similar socioeconomic standing (Mostafa, 2010). Thus, students of a similar socioeconomic background enroll in schools with others of similar backgrounds. Thus, the research indicates that the background of the student plays a significant role in influencing academic achievement.

However, achievement scores are not solely a function of the students' background, but instead represent a confluence of other variables, of which the educational system is paramount. Therefore, student scores also fail to reveal the influence of institutional effects on student achievement. Disparities in educational funding from community to community results in reduced resources for less privileged students (Mostafa, 2010). Furthermore, those disparities are compounded by difficulties recruiting and retaining effective teachers (Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Tillman, 2005), thereby increasing the proportion of minority and low-income students exposed to less effective teachers (Murphy, DeArmond &Guin, 2003; National Science Foundation, 2012a).

Research has consistently indicated the importance of effective teachers in improving student achievement for all students (Allen, Gregory, Mikami, Lun, Hamre & Pianta, 2013; Konstantopoulos & Chung, 2011). Studies have indicated that lower teacher expectations often lead to weakened academic performance in students

(Gregory & Weinstein, 2008; Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Rumberger & Palardy, 2005; Strand, 2012. For example, research by Konstantopoulos and Chung (2011) demonstrated that teachers have a significant influence on student achievement, particularly in reading and science.

This effect is cumulative. Having effective teachers year after year produced larger gains in achievement for all students, an impact which was particularly pronounced in schools with high minority populations. According to the study's authors, a single effective teacher increased student achievement by approximately one tenth of a standard deviation (Konstantopoulos & Chung, 2011). Therefore, considering that effective teachers are more likely to be concentrated in schools with students of high socioeconomic standing (Bankston III & Caldas, 1998; Caldas & Bankston III, 1997; Condron & Roscigno, 2003; Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Rumberger & Palardy, 2005; Ryabov &Van Hook, 2007), the experience of schooling differs remarkably based on one's socioeconomic background.

Everson and Millsap (2004) revealed that those differences between schools matter as those differences are responsible for producing and promulgating differences in student achievement as a factor of school size, the ethnic and racial composition of schools, and the proportion of students in poverty. Okpala, Smith, Jones, and Ellis (2000) and Perry and McConney (2010) also revealed that schools with a high composition of students of low socioeconomic status (SES) have lower levels of achievement in math and reading. In fact, research has indicated that when individual family backgrounds are controlled, the socioeconomic composition of schools remains a strong predictor of student achievement (Kahlenberg, 2001; Perry & McConney, 2010).

Teacher expectations of student performance and the type of curriculum provided are also influenced by the socioeconomic composition of schools (Gregory & Weinstein, 2008; Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Rumberger & Palardy, 2005; Strand, 2012). However, a highly cohesive school characterized by a supportive and inviting environment was found to mitigate the educational ills that are pervasive in large, urban, primarily minority schools (Stewart, 2007). Thus, research on factors affecting student achievement, as reflected by standardized test scores, is clear; student achievement is a multifaceted outcome of an amalgamation of both student social background and systemic/institutional effects. The results of these differences are readily apparent in scores obtained from standardized measures that elucidate how educational attainment is a product of variables, most of which are beyond the control of the student.

For example, an analysis of students results on the National Assessment of Educational Progress (NAEP) revealed that minority students attained significantly lower reported scores than their White or Asian/Pacific Islander counterparts while female students scored lower than their male counterparts and students from lower-income households scored significantly lower than their peers from more affluent households (National Center for Education Statistics, 2012). Science proficiency particularly demonstrates this disparity in academic achievement (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008; Florida Department of Education, 2012b).Characteristics such as ethnicity, socioeconomic status, and gender have a significant influence on achievement scores in science.

While these student characteristics have been deemed influential on student achievement in science, current research has failed to investigate their impact on biology at the secondary level. Some studies have utilized biological content on assessments of scientific knowledge (along with questions from other branches of science), but there is a dearth of research about the extent that socioeconomic and other factors influence on biology achievement. The limited research available is dated, focuses on affective or metacognitive factors, only evaluates online instruction, or concentrates specifically on particular instructional strategies.

However, researchers have largely ignored investigating these student characteristics possibly due to the assumption that since these student characteristics are influential in other branches of science and academic domains, they probably are influential in biology achievement. Therefore, a formal investigation would be futile and redundant. However, recent developments may have a significant impact on biology student achievement and are worth investigating.

First, there has been an increase in the implementation of state-mandated high school biology assessments nationwide (Blazer, 2012), thus prompting the allocation of more fiscal and personnel resources to biology instruction. Other secondary science subjects have not received the same degree of increase in resources, so it could have a significant impact specifically on biology. Second, there has been an increase in the number of secondary minority and female science teachers (Wood, 2002), which may potentially increase female and minority students' identification with and interest in the sciences.

Other significant developments include: the increased literacy rate of minority and low socioeconomic students (National Center for Education Statistics, 2011); increased allocation of fiscal resources and personnel to improve literacy outcomes (Golden, 2003; Winstead, 2011); the majority of high school students nationwide who take biology by their high school graduation (Blank, Langesen, & Petermann, 2007); an increase in the number of students taking advanced coursework in biology (Blank, Langesen, & Petermann, 2007; National Science Board, 2012); an increase in the number of minorities and low socioeconomic students enrolling in advanced science courses (National Science Board, 2012); the increased number of female secondary students opting for more advanced coursework in science (Moore & Slate, 2008); the increased number of certified biology teachers (Blank, Langesen, & Petermann, 2007); the increased number of science teachers with advanced degrees and full certification (National Science Board, 2012); the impact of the majority (93%) of biology teachers who teach in-field (National Science Board, 2012); the fact that more than half of secondary biology teachers who are female with 54% possessing a master's degree and more than 10 years teaching experience (Wood, 2002); a higher proportion of females undertaking biological collegiate coursework (National Science Foundation , 2011; Sonnert & Fox, 2012); a higher representation of minorities obtaining bachelor's and master's degrees in science since 1989 (National Science Foundation, Division of Science Resource Statistics, 2011); more women than men attaining undergraduate scientific degrees (National Science Foundation, 2012b); and greater numbers of women earning doctoral degrees in the sciences since 2004 (National Science Foundation, 2011). These developments undoubtedly have some impact on student

achievement in biology, but the question remains of how influential socioeconomic and other factors are on student achievement. This is especially noteworthy due to the increased emphasis on biology achievement nationwide as measured by end-of-course tests instead of the reliance on comprehensive science assessments. This study was developed to address this gap in the empirical evidence base.

It also seeks to address the paucity of research on the importance of reading proficiency for biology achievement, particularly at the secondary level. While numerous studies have been conducted on the importance of reading for comprehensive science assessments, only one to date has specifically addressed the role of reading on biology achievement. However, this study was conducted at the tertiary level, thereby underscoring the importance of secondary-level specific studies. Consequently, researchers are relying heavily on results obtained from comprehensive science assessments to make inferences on the influence of reading on biology achievement.

While such inferences are not erroneous, each scientific discipline varies in its literacy demands, and the magnitude of reading's influence on biology achievement at the secondary level has yet to be determined. The disparities in literacy demands and the degree of reading influence are particularly pertinent considering the need to develop appropriate professional development opportunities for teachers as well as the significant ramifications of assessment scores on school accountability. Thus, the intent of this study is to address the gaps in extant literature so that instructional leaders, administrators, teachers, stakeholders, and policymakers can make more informed decisions.

End-of-Course Assessments

In the past most states utilized comprehensive assessments to satisfy school accountability requirements. Recently, more states have shifted from utilizing comprehensive exit exams that assess multiple subjects on a singular test, to end-of-course (EOC) assessments that measure student mastery of course content (Center on Educational Policy, 2008; McIntosh, 2012). EOC assessments are comparable to the final exams that are completed at the end of a course, but they are statewide standardized tests (Blazer, 2012; McIntosh, 2012). Florida, for example, has moved from requiring 11th grade students to complete the Florida Comprehensive Assessment Test (FCAT) Science, which covers various science course content from grades 9 through 11, to mandating that all students enrolled in biology I or an equivalent course complete the Biology EOC assessment in order to receive the course credit required for graduation. Twenty-two states administered EOC exams to students in associated courses as of spring 2012 (Blazer, 2012). Over the next 10 years, that number is expected to increase to 26 states (Blazer, 2012). The number of EOC assessments administered per state varies from 1 to 16, with Florida electing to implement five (Blazer, 2012; McIntosh, 2012).

The nationwide popularity of EOC assessments is not without its detractors. Although the EOC movement is designed to connect course curriculum standards more closely with high school assessments so that students are increasingly prepared for postsecondary education or entry into the workforce, critics are concerned about the costs associated with statewide implementation and about the possibility of excessive testing in schools (Blazer, 2012). Other opponents include teachers' unions and advocates of students with special needs (McIntosh, 2012) who cite the

disproportionately negative effects of testing on subsets of the student population and the increasing consequential inequality between groups.

For example, research has revealed that due to the increased accountability measures, Hispanics and African Americans have demonstrated smaller academic gains relative to White students, thus widening the achievement gap (Hanushek & Raymond, 2005). Research has also indicated that minorities pass the standardized tests required for graduation at significantly lower rates. According to Borg, Plumlee, and Stranahan (2007), "an average White student has a 65% probability of passing the FCAT graduation requirement on the first try as compared to a 34% probability for an identical African American student... An average Hispanic student has a 54% probability of passing on the first try—11 percentage points lower than an identical White student" (p. 712). Furthermore, research has also exposed that standardized testing in Florida has a disproportionate effect on African American and Hispanic students from low socioeconomic, less educated, and high mobility households who are less likely to satisfy graduation requirements than their higher socioeconomic, White, suburban peers living in educated households (Borg, Plumlee, & Stranahan, 2007).

Failure to pass standardized tests can have lasting consequences. Minority and low-income students who barely failed New Jersey's exit exam were more likely to drop out of high school (especially those students who failed on their first attempt) than students who barely passed (Ou, 2009). Papay, Murnane, and Willett (2010) found similar results. Low income students from urban environments were more susceptible, on average, than their wealthier, suburban peers, to the effects of failing an exit exam. They are more likely to drop out of high school if they barely fail the assessment on their

first attempt than their peers who barely passed. McIntosh (2012) also revealed that high stakes testing affects 7 of every 10 public school students, particularly students of color (71% of African American and 85% of Hispanic students), English language learners (83%), and those of low socioeconomic status (71%) that reside in the 26 states that administer or will implement exit exams. Thus these assessments are more commonplace in states that contain higher proportions of economically disadvantaged and minority students (Center on Educational Policy, 2008), so the consequences for these demographic groups are magnified.

In Texas, EOC assessments serve as a graduation requirement for students enrolled in grade 9 during the 2011-2012 school year (Blazer, 2012). According to Blazer (2012), students would be required to pass four EOC assessments: English, mathematics, science, and social studies. In response to the adoption of this graduation requirement, superintendents from various districts testified at a Texas House Public Education Committee meeting that dropout rates might increase because students would believe that failing two or three EOC exams would make it impossible to catch up. In fact, 75% of students who failed their EOC assessments in spring 2012 were considered at risk of dropping out of school, according to the Texas superintendents (Blazer, 2012).

Therefore, the increased focus on stronger accountability measured through standardized tests has unintentionally increased gaps in achievement and opportunity. Rather than narrowing the chasms between students of varying ethnicities and socioeconomic in student achievement and graduation rates (Hanushek & Raymond, 2005), they are being exacerbated. Since graduation from high school is associated

with varied positive life outcomes, educational policymakers must understand the impact of high-stakes testing on high school completion (Papay, Murnane, & Willett, 2010).

Despite the legitimacy of these concerns, advocates of the EOC assessments have remained vocal. Higher education officials have supported the implementation of these assessments, citing the considerable time and costs associated with student remediation of incoming freshman who are unprepared for collegiate-level work (McIntosh, 2012). The business community supports assessments, contending that they help overcome the lack of basic skills that some recent graduates display in entry-level positions (Center on Educational Policy, 2008; McIntosh, 2012). Other advocates of EOC assessments believe that these assessments are better predictors of success at the collegiate level and readiness for work than grades obtained in the classroom or comprehensive exit exams (Blazer, 2012).

Research has substantiated some of these claims. According to Hanushek and Raymond (2005), research has indicated that since the introduction of accountability systems during the 1990s, student achievement, as measured by state achievement growth data from the National Assessment of Educational Progress (NAEP), has increased. This is particularly true for states that implemented consequential accountability earlier, as they exhibited more rapid gains in NAEP performance, particularly among Hispanic students (Hanushek & Raymond, 2005). Carnoy and Loeb (2002) also revealed that in states with strong state accountability, students averaged significantly greater gains across all groups on the 8th-grade math NAEP and also

achieved at higher levels of math than students in states with little or no accountability measures.

Reasons cited for implementation of the EOC

The purpose of implementing the EOC assessment varies according to state (Blazer, 2012). McIntosh (2012) suggests that some states adopt EOC exams in subjects other than English and math in order to evaluate teachers. As EOC assessments allow for easier associations between test scores and specific teachers, they serve as better gauges of teacher performance than other assessments (McIntosh, 2012). McIntosh (2012) also hypothesized that EOC assessments afford greater flexibility at the secondary level since they can test a wide variety of course content that is a function of varied course selection by students. Hence, the connection between curriculum and instruction is improved compared to other kinds of assessments (McIntosh, 2012).

In any case, states have reported that they use the assessment to determine student mastery of state standards or as a tool to identify students who are at risk of dropping out of school (Blazer, 2012; McIntosh, 2012). Others use the EOC results to calculate final course grades and or to require a passing score as a graduation requirement (Blazer, 2012; McIntosh, 2012; Zinth, 2012). In Florida, EOC assessments are criterion-referenced tests that measure student proficiency with the Next Generation Sunshine State Standards for specific courses and passing this assessment is a graduation requirement (Blazer, 2012).

The implementation of the end-of-course assessments

During the 2011- 2012 school year 25 states administered comprehensive or EOC exams on some variation of English language content and 23 assessed math

(McIntosh, 2012), but currently only 21 states are administering EOC assessments in biology (Blazer, 2012). The implementation of EOC assessments in science, specifically biology, is relatively new as past assessments focused primarily on reading and mathematics. This addition has been welcomed within the scientific community. Advocates of science education, such as Bybee (2010) campaigned for the equal treatment of science in education reform by pointing out how the No Child Left Behind act inadvertently reduced or eliminated science from some school programs (particularly at the elementary level) by failing to include science test scores as a significant part of computing adequate yearly progress.

Student Performance on Standardized Science Tests

Such shifts in priorities demonstrate how Florida diverted resources and curriculum emphasis as a result of No Child Left Behind mandates regarding standardized test performance on mathematics and reading. Research has indicated that educators have diverted resources or increased funding to focus on reading and mathematics while neglecting non-tested subjects and reducing funding for gifted students (Darling-Hammond, 2006; Doppen, 2007, Golden, 2003, Marchant, 2004, Winstead, 2011). Research on the effects of school decisions to focus primarily on assessed subjects has also revealed: the inclusion of reading and mathematics in the subject-area classroom (Velde Pederson, 2007; Winstead, 2011); a reduction in instructional time for other subjects (Velde Pederson, 2007; (Winstead, 2011) a decline in resource availability for other subject areas; the allocation of additional periods of reading instruction to students who fail to attain proficiency on standardized tests; and decreased professional development in subjects other than reading and math (Winstead, 2011).

Thus, the extent to which those curriculum changes and diverted resources has affected biology achievement has not been documented. Such changes could potentially result in increased biology achievement due to improved student reading comprehension or a reduction in biology achievement due to loss of instructional time, resources, and professional development opportunities for teachers. Further research is needed to explore this outcome.

The need for higher quality science instruction is readily apparent when reviewing U.S. student performance on various international and national assessments. One such assessment, the Trends in International Mathematics and Science Study (TIMSS), measures science knowledge and skills of fourth and eighth-graders internationally. TIMSS results in 2007 indicated that U.S. fourth graders ranked eighth out of thirty-five countries in science achievement and our eighth-graders eleventh out of forty-seven countries in 2007 (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008). U.S. students also performed in higher percentages in each of the benchmarks regarding proficiency levels at both grade levels; thus there were proportionally more U.S. students who performed at the advanced level than the international mean and concurrently more U.S. students underperforming compared to the international mean (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008). When U.S. fourth- and eighth-grade student performance was compared over a twelve-year period (using data from 1995, 2003, and 2007), results indicated no significant change in student performance at both grade levels (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008).

Another international assessment, the Program for International Student Assessment (PISA) measures scientific literacy of 15-year olds in select countries. U.S. students ranked number 23 among both 65 Organisation for Economic Co-operation and Development (OECD) and Non-OECD countries in 2009 (Fleischman, Hopstock, Pelczar, & Shelley, 2010). Only 29% of U.S. students scored at the literacy level associated with high order tasks such as "select[ing] and integrat[ing] explanations from different disciplines of science and technology" and "link[ing] those explanations directly to...life situations" (Fleischman, Hopstock, Pelczar, & Shelley, 2010, p. 26). Eighteen percent of U.S. students scored at a level considered below baseline level of proficiency in which "students begin to demonstrate the science competencies that will enable them to participate effectively and productively in life situations related to science and technology" (Fleischman, Hopstock, Pelczar, & Shelley, 2010, p. 26). Nevertheless, the average U.S. science literacy score increased between 2006 and 2009, although 71% of U.S. students in 2009 still fall within the category of low to moderate levels of proficiency in science (Fleischman, Hopstock, Pelczar, & Shelley, 2010).

When the National Assessment of Educational Progress (NAEP) assessed the scientific knowledge of U.S. elementary and secondary students, , 65% of eighth-graders performed at or above basic level, 32% at or above proficient, and 2% at or above advanced in 2011 (National Center for Education Statistics, 2012). When compared with results obtained in 2009, the percentage of students at or above basic and proficient levels increased by 2 points, although there was no significant change in the percentage of students performing at the advanced level (National Center for Education Statistics, 2012).

Additionally, an examination of student results on the ACT indicates that the average ACT score in science has remained relatively stable since 2006 (ACT, 2010). Furthermore, only 29% of ACT-tested high school graduates in 2010 met the College Readiness Benchmarks score in science. Since 2006, the percentage of students who met the College Readiness Benchmarks in Science remained relatively stable with a 2% increase in 2010 from 2006 (ACT, 2010).

When the results of all these assessments are collectively examined, one can readily identify the consistent underperformance of U.S. students in science. Although there are indicators of improvement, significant inroads need to be made in improving student achievement nationally.

An examination of Florida's results on the NAEP reveals that although the state's average 8th-grade science scores improved by 3 points from 2009 to 2011, the average score was below the national average by 3 points (National Center for Education Statistics, 2012). Furthermore, 72% of Florida's students were categorized as below basic or basic in science proficiency, with only 1% scoring at the advanced level (National Center for Education Statistics, 2012). Thus, the statistics paint a gloomy picture of student preparation for collegiate and science careers as well as overall improvement in the future general public's understanding of science.

These statistics also allude to the increased need to identify the student characteristics that play a significant role in affecting student achievement in science. Although the aforementioned assessments were not tied to school accountability indicators, they nevertheless suggest that a significant proportion of Florida's students may not achieve the passing score required for graduation – assuming the student

performance on Florida's Biology end-of-course is relatively similar. Thus, it is essential to identify those students at risk of failing to meet Florida Department of Education's accepted standard of achievement in biology.

Therefore, although Florida’s education policy centers on improving student achievement by increasing the rigor, skills, and core content knowledge by the implementation of increasing number of standards and assessments, this assertion is, based on the research, highly improbable given the confluence of variables that impact on student achievement. However, since more changes can be made to teaching rather than altering students' personal characteristics or family circumstances; educational policy discussions will continue to remain focused on what is within a school's control. Therefore an investigation of the student characteristics that influence biology achievement is a timely one. It is essential that education personnel are aware of the role of select student characteristics, such as reading proficiency, ethnicity, socioeconomic status, and gender; that research has consistently shown to influence student achievement, and their impact on student achievement on the Biology I end-of-course assessment.

Statement of the Problem

Due to the increasing high-stakes associated with student performance on the Biology I EOC assessments, schools need to be able to predict and determine the influence of student characteristics on student performance prior to the administration of the test. Although a significant number of studies in science, technology, engineering, and math (STEM) have produced viable answers, there is a paucity of research specifically on the predictors and differences in student achievement in biology specifically. Thus stakeholders are limited when making decisions about biology

education. By exploring four variables- reading proficiency, ethnicity, socioeconomic status, and gender- the researcher can address the gap in extant literature and stakeholders will be able to potentially make instructional changes that can improve student performance on the assessment and reduce achievement gaps.

Such changes could have a profound impact on the success of individual schools. Student outcomes affect a school's grade which determines whether it is considered high-performing and receives additional funding by the federal government. If it is labeled underperforming, it could eventually be closed or taken over. t As the Biology end-of-course assessments are tied to graduation sanctions, the graduation rate would also be affected by changes in biology achievement. Graduation rate is itself another accountability indicator which has a significant impact on schools. Moreover, the performance of students also serves as the primary basis for teacher evaluations (The Florida Legislature, 2010) and thus has significant ramifications on the recruitment, retention, and morale of highly qualified and skilled teachers, especially those teaching in high-poverty schools.

Research Questions

1. Is there a significant difference in student performance on the end-of-course assessment in biology associated with reading proficiency? After adjusting for student gender, race/ethnicity, and socioeconomic status, is there a significant difference in student performance on the end-of -course assessment in biology associated with reading proficiency?

2. Is there a significant difference in student performance on the end-of-course assessment in biology associated with race/ethnicity? After adjusting for reading proficiency, student gender, and socioeconomic status, is there a significant difference in student performance on the end-of-course assessment in biology associated with race/ethnicity?

3. Is there a significant difference in student performance on the end-of-course assessment in biology associated with socioeconomic status? After adjusting for reading proficiency, race/ethnicity, and student gender, is there a significant

difference in student performance on the end-of-course assessment in biology associated with socioeconomic status?

4. Is there a significant difference in student performance on the end-of-course assessment in biology associated with gender? After adjusting for reading proficiency, race/ethnicity, and socioeconomic status, is there a significant difference in student performance on the end-of -course assessment in biology associated with gender?

5. What are the best predictors of student performance on the biology end-of-course assessment?

Statement of Hypothesis

It was hypothesized that there will be a significant difference in student performance on the end-of-course assessment in biology associated with reading proficiency. It is hypothesized that students who are more proficient readers will score significantly higher than their less proficient peers on the Biology I EOC assessment. Furthermore, more proficient students will score significantly higher that their less proficient peers on the Biology I EOC assessment after adjusting for student gender, race/ethnicity, and socioeconomic status.

It was hypothesized that there will be a significant difference between the race/ethnicity of a student and their performance on the end-of-course assessment in biology. It is hypothesized that students from Caucasian backgrounds will score significantly higher than those from other ethnic backgrounds. It is also hypothesized that students from Caucasian backgrounds will score significantly higher than those from other ethnic backgrounds after adjusting for reading proficiency, gender, and socioeconomic background.

It was hypothesized that there will be a significant difference in student performance on the end-of-course assessment in biology associated with socioeconomic status. It is hypothesized that students from higher socioeconomic

backgrounds will score significantly higher than those from lower socioeconomic backgrounds. It is also hypothesized that students from higher socioeconomic backgrounds will score significantly higher than those from lower socioeconomic backgrounds after adjusting for student gender, race/ethnicity, and reading proficiency.

It was hypothesized that there will be a significant difference in student performance on the end-of-course assessment in biology associated with gender. It is hypothesized that males will score significantly higher than females on the Biology I end-of-course assessment. It is hypothesized that males will score significantly higher on the end-of-course assessment in biology after adjusting for reading proficiency, race/ethnicity, and socioeconomic background.

It was hypothesized that reading proficiency, ethnicity, and socioeconomic status will serve as significant predictors of student performance on the Biology end-of-course assessments.

Significance of the Study

The purpose of this study is two-fold. First, it seeks to determine how student performance is influenced by four variables--reading proficiency, ethnicity, socioeconomic status, and gender--and examines their impact on the results of the Florida Biology I EOC assessment at a Title 1 school. The study is intended as an initial source of feedback to instructional leaders, teachers, and administrators regarding the status of presumed achievement gaps based on student characteristics and it may prompt further discussion on institutional practices to alleviate those gaps. Second, it aims to predict student performance in biology. Thus, it aims to address the gaps in empirical studies regarding the student characteristics that impact biology achievement.

This study's findings will contribute to research on student achievement in biology and the effect of end-of-course assessments on graduation rates, school accountability indicators, professional development, and teacher evaluations. Previous research findings on many of these variables have been limited or nonexistent, and therefore this study can contribute to the existing literature. These findings may also be useful within the state of Florida and other states that have implemented or plan on implementing end-of-course science assessments instead of minimum competency based graduation tests.

The majority of existing literature on student achievement in biology has focused on a single predictor. This study is unique in examining four variables simultaneously, each of which has been individually identified as very influential on student achievement in science. By including these four variables simultaneously, the model is capable of determining their cumulative effect on student achievement rather than focusing on and overemphasizing one variable.

This study will analyze individual student scores while simultaneously controlling for other variables through adjusting. Undertaking this task is important as influential variables due to diverse student backgrounds (reading proficiency, ethnicity, socioeconomic status, and gender) intersect with each other. Thus the results of the study can be utilized as a source of insight in designing a comprehensive plan to address the achievement gap in science at the school-site level.

Unlike other assessments of student performance in science—such as the National Assessment of Educational Progress (NAEP), Program for International Student Assessment (PISA), the Third International Mathematics and Science Study

(TIMMS), and FCAT Science— the Florida Biology I end-of-course assessment serves as the basis of not only school accountability, but also that of students and teachers. The previously adopted FCAT Science assessments were solely utilized to evaluate schools. These tests are not utilized by teachers to assign course grades and scores are not recorded on high school transcripts, so there is no motivation or incentive on the student's part to do well on such tests (unless they are concerned about the statewide prestige of their school and potential funding opportunities from governmental agencies) (Stansfield, 2011).

However, this does not invalidate the usefulness of the other aforementioned assessments in providing data that clearly indicates discrepancies in student performance as a factor of ethnicity, socioeconomic standing, and gender at the state, national, and international levels. Thus, using data derived from the Florida Biology EOC assessment will serve as a useful tool for comparison with previous national and international findings to determine whether similar results will be replicated.

Other studies have not made use of the state's reading assessment (FCAT Reading and FCAT 2.0 Reading) as the measure of reading proficiency. Thus for administrators and teachers within the state of Florida, this data is more familiar and readily accessible. Depending on the study's outcome, this source of data can become more significant as a potential means of making initial decisions regarding student achievement in biology. Furthermore, the future of educational policy decisions by site administrators and district and state personnel may be potentially affected by outcomes of the study.

Finally, these findings have greater implications for the future of this nation and its ability to educate all students in a challenging manner that enables each student to obtain the knowledge and skills needed to become scientifically literate. The results can serve as one of many indicators to gauge trends in scientific proficiency, collegiate preparation, and the status of achievement gaps based on ethnicity, socioeconomic status, and gender. Thus, the results of this study can be used to determine instructional deficiencies, instructional practices, or systemic issues that contribute to an achievement gap so that stakeholders can potentially rectify such issues. Educators cannot afford to be cavalier about what student achievement data reveals about the degree of preparedness of our students for a more advanced technological society. It has dire consequences for our ability to compete globally in the future, and influences the degree to which an informed citizenry can contribute to the democratic process. There are additional implications regarding national immigration and security policies.

Limitations

One of the study's most important limitations is part of the nature of a causal-comparative study. Causal-comparative studies aim to determine the cause for existing differences in the behavior or status of groups or individuals (Gay, Mills, & Airasian, 2009). Because the cause is preexisting, the researcher has no control over the conditions of this *ex post facto* study and cannot draw cause-effect conclusions with any degree of confidence due to the lack of randomization and manipulation. Thus, the interpretation of findings in this study requires considerable caution.

Additionally, the study's use of convenience sampling means that the sample size is limited and not representative of either the demographic composition or the academic proficiency of students in the state of Florida. Furthermore, due to the use of

data obtained from a Title 1 school, the effect of school context on student achievement is not addressed by the study.

CHAPTER 2
REVIEW OF THE LITERATURE

A study on variables that impact student achievement on the Florida Biology end-of-course assessment would not be complete without a review of related literature. Thus the following discussion will present an overview of research related to this study. The following subsections that comprise this chapter include: (a) Standardized Testing, (b) The Importance of Science Education, (c) Reading and Science Achievement, (d) Ethnicity and Science Achievement, (e) Socioeconomic Status and Science Achievement, and (f) Gender and Science Achievement. For purposes of this study, the literature review subsection on ethnicity will focus primarily on African American, Hispanic, and White populations as the researcher has elected to discuss the major demographic groups that are representative of the school under investigation.

Standardized Testing

According to Haladyna, Nolen and Haas (1991), standardized tests are universally and uncritically accepted by the public and many educators as a valid measure of educational accomplishment “and an increased perceived need to evaluate education at virtually all units of analysis (i.e., individuals, classes, schools, school districts, states, and even the nation)” (p. 3). These scores are considered “high stakes” because tangible consequences depend on them and they are used to: rank schools and their districts to bemoan the failure of education in newspapers, determine merit pay and make other personnel decisions by school district personnel, used to rate neighborhoods in terms of its “quality” by real estate agents, and rank states and their effectiveness on state and national levels; to name a few (Haladyna, Nolen & Haas, 1991; Madaus & Russell, 2010; Wiliam, 2010). Thus, these high stakes tests are

probably going to remain a thing of the future as long as test scores remain the single most important index of educational effectiveness (Haladyna, Nolen & Haas, 1991).

Although these high stakes assessments provide an insight into the academic achievement of students nationally and at the state level, no single standardized achievement test represents a complete mapping of the content of school achievement domain (Haladyna, Nolen & Haas, 1991; Madaus & Russell, 2010; Wiliam, 2010). Often test results are used without considering the complexity of achievement and its causes (Madaus & Russell, 2010). For example, attributing the level of achievement based on test scores to the influence of a single teacher, school, or school district erroneously oversimplifies the nature of the scores, thus contributing to the loathing of teachers and administrators toward standardized testing (Haladyna, Nolen & Haas, 1991). Furthermore, student achievement is a function of a variety of factors, of which only some are under the influence of schools (Haladyna, Nolen & Haas, 1991; Madaus & Russell, 2010; Qiu & Wu, 2011). Schools can influence the quality and quantity of instruction, motivation, and the learning environment but they have little to no effect on family and home environment, maturity, and mental ability (Haladyna, Nolen & Haas, 1991).

Test Pollution

Typically in response to the high-stakes nature of standardized testing, test score pollution occurs. Test score pollution refers to factors affecting the truthfulness of a test score interpretation and is very pervasive in America (Haladyna, Nolen & Haas, 1991). The main sources of test pollution are: the way schools and its personnel prepare students for tests, test administration activities or conditions, and exogenous factors that are beyond the control of schools and school personnel (Haladyna, Nolen & Haas,

1991). Polluting practices include: teaching test-taking skills, preparing teaching objectives and developing a curriculum to match the test, using commercial materials specifically designed to improve test performance, presenting test items before the test, and interfering with response (e.g. giving hints or answers to students or altering response sheets). Other known conditions include students' anxiety, stress, fatigue, and motivation which all impede student performance.

In regards to exogenous factors, the reporting of test scores without acknowledging the influence of the family, family mobility, economic environment, proficiency with the English language, and other such factors results in the depiction of invalid inferences from the test scores (Haladyna, Nolen & Haas, 1991; Madaus & Russell, 2010). In fact, past research has documented the significant impact of student family background on test scores, in which single-parent backgrounds had a negative effect (Bankston III & Caldas, 1998; Pong, Dronkers, & Hampden-Thompson, 2003; Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Qiu & Wu, 2011; Rathbun & West, 2004), and increasing income and education levels were associated with positive effects (Bankston III & Caldas, 1998; Education Information and Accountability Services, 2009; Liu & Lu, 2008; Qiu & Wu, 2011; Rathbun & West, 2004; Woods, Kurtz-Costes, & Rowley, 2005). Therefore, the socioeconomic status of the student affects educational opportunities and accessibility to a variety of resources while simultaneously the home environment affects the degree of student aspiration in learning and interest in school (Qiu & Wu, 2011).

Research has also indicated that the validity of standardized tests is always a source of concern as although they are carefully designed to measure predetermined

constructs, the readability of test items is often considered as a threat to the validity of the test (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1999). Therefore, according to the *Standards for Educational and Psychological Testing* (American Educational Research Association, et al., 1999), the possibility that the test is assessing reading proficiency and reading comprehension in addition to its intended concepts, is increased if test items are not examined for readability and comprehension to the test takers. It is therefore recommended that "in testing applications where the level of linguistic reading ability is not part of the construct of interest, the linguistic or reading demands of the test should be kept to the minimum necessary for valid assessment of the intended construct" (American Educational Research Association et al., 1999, p. 82). Thus, the length of sentences, selected vocabulary, and the direct nature of questioning should be relatively easy in communicating the construct so that student knowledge of the construct, not reading proficiency, is assessed (Visone, 2009).

Florida Biology I End-Of-Course (EOC) Assessment

According to Goodwin, Englert, and Cicchinelli (2003), the original purpose for outcome-oriented accountability systems is based on the assumption of improving schools by increasing academic performance for all students. This premise serves as the basis for the implementation of the Biology I end-of-course (EOC) assessment as part of Florida's Next Generation Strategic Plan (Florida Department of Education, 2005a). This Next Generation Strategic Plan operates under the vision that "Florida will have an efficient world-class education system that engages and prepares all students to be globally competitive for college and careers" (Florida Department of Education, 2005b) and which would be measured by EOC assessments. According to the Florida

Legislature (2012), the Florida Statutes state that, "the primary purpose of the student assessment program are to provide information needed to improve the public schools by enhancing the learning gains of all students and to inform parents of the educational progress of their public school children."

The program was designed to satisfy six conditions: 1) to assess the learning aims of each student annually in achieving the Sunshine State Standards appropriate to the grade level of the student, 2) to provide data for school accountability and recognition so that schools can make decisions, 3) to determine the educational strengths and needs of students as well as their readiness to be promoted to the next grade level or to graduate from high school with either a standard or special high school diploma, 4) to determine how well educational goals and curricular standards are achieved at the school, district, and state level, 5) to provide information that can be used in the evaluation and development of educational programs and policies, and 6) to determine student performance in Florida against that of other students nationally (The Florida Legislature, 2012).

One such assessment, the Biology I end-of-course assessment; was administered for the first time in May 2012 to students enrolled in biology I or an equivalent course to determine student proficiency with the Next Generation Sunshine State Standards (NGSSS) and expectations for student learning outlined in the course descriptions (Florida Department of Education, 2005a). Student performance during this academic year on the Biology I EOC would constitute 30% of the student's final course grade (The Florida Legislature, 2012). In addition to this and other student performance indicators, student performance contributed to the school's overall grade, and served as

the chief criteria for biology teacher evaluations (The Florida Legislature, 2010; 2012). For the first administration of the assessment, student achievement was categorized in achievement levels represented by thirds based on comparisons with other students in Florida that range from 1 (lowest third) to 3 (highest third) (Florida Department of Education, 2012i).

During the 2012-2013 school year, all students entering grade 9 would be required to earn a passing score on the EOC assessment in biology to earn course credit (The Florida Legislature, 2012). This passing score indicated by obtaining a minimum scale score in Achievement Level 3, will be used to fulfill one of the requirements for high school graduation (Florida Department of Education, 2012f). Future test administrations starting during the 2012-2013 academic year will also have a similar effect on the overall school grade and teacher evaluations (The Florida Legislature, 2012). Thus this new current policy of science accountability will not only present a new set of challenges: one of which is due to the 9th grade reading level in which the test is written and its incorporation of short reading passages (Florida Department of Education, 2011c), and the historical differing levels of achievement as an outcome of the diversity of the student body; but also opportunities in science education.

One such opportunity is the reduction in the achievement gap between subgroups of the student population (Penfield & Lee, 2010). Additionally, test-based accountability in science forces schools, districts, and states to ensure that the educational needs of all students, especially those in historically low-performing subgroups; are addressed (Penfield & Lee, 2010). Thus, this commonly overlooked

subset of the school population will be able to gain access to educational opportunities and the associated potential for careers in science, thus reducing inequalities already present within the educational sector. Last, science will be integrated within the school curriculum after traditionally being ignored in order to address the urgency associated with basic literacy and numeracy (Doppen, 2007; Penfield & Lee, 2010; Winstead, 2011). Thus, additional resources and instructional time for science may now be potentially allocated by districts and schools (Penfield & Lee, 2010).

Results of first administration of the Biology I EOC

As previously discussed, student achievement levels for the May 2012 assessment administration was categorized in thirds in which student T-scores were compared with other students in Florida (Florida Department of Education, 2012i). Scores ranged from 1 or lowest third to 3 or highest third. Of the 190,344 students who took the Biology EOC assessment in May 2012, 35% scored in the top third while 31% scored in the middle third and 34% the lowest third (Florida Department of Education, 2012a). Using scores represented on the T-scale ranging from 20-80, the statewide mean was approximately 49 (Florida Department of Education, 2012a).

Results of the assessment revealed that only 59% of students would have obtained a passing score on the assessment, of which 11% would score at mastery level 5, 11% at the above satisfactory level 4, 37% at the satisfactory level 3, 27% at the below satisfactory level 2, and 14% at the inadequate level of 1 (Florida Department of Education, 2012m). Furthermore, as student grade level increased (beginning with the 7th grade and ending with 12th grade), the mean scale score for each grade level decreased. Thus 7th grade students had the highest mean scale score of 60 whereas their 11th and 12th grade counterparts obtained a mean scale score of 43 (Florida

Department of Education, 2012a). Additionally, as the student grade level increased, the percentage of students scoring in the top third decreased significantly (Florida Department of Education, 2012a). For example, at least 50% of students enrolled in grades 7,8, and 9 scored in the top third whereas less than 25% of 10th, 11th, and 12th grade students earned a comparable score (Florida Department of Education, 2012a). Furthermore, students enrolled in higher grade levels were disproportionately represented in the lowest third than their peers enrolled in lower grades (Florida Department of Education, 2012a).

When the test results were further analyzed, certain patterns became apparent regarding ethnicity, socioeconomic status, and gender. According to Florida Department of Education (2012b), a greater percentage of Asian and White students scored in the top third than Black and Hispanic/Latino students. In fact, 58% of Asian and 45% of White students scored in the top third compared to 17% of Black students and 29% of Hispanic/Latino. Black and Hispanic students were also disproportionately represented in the lowest third. 53% of Black and 39% of Hispanic students comprised the lowest third compared with 23% of White students (Florida Department of Education, 2012b).

When data was disaggregated regarding socioeconomic status, other trends surfaced. For each grade level, higher percentage of students who were ineligible for free or reduced lunch scored in the top third than eligible students (Florida Department of Education, 2012b). In fact, 23% of those eligible for free and reduced-price school lunch scored in the top third compared to 48% of those ineligible. Additionally, 45% of those eligible for free and reduced-price lunch scored in the lowest third compared to 22% of those ineligible (Florida Department of Education, 2012b).

In terms of gender, simple patterns as well as the intersection of gender and ethnicity surfaced. For each grade level, higher percentage of males scored in the top third than females (Florida Department of Education, 2012b). When ethnicity was analyzed with gender, data revealed that only 16% of Black males compared to 47% of White and 31% of Hispanic/Latino males scored in the top third. Of those scoring at the lowest third, the majority were Black, specifically 55% of Black males compared to 39% of Hispanic/Latino, and 23% of White males (Florida Department of Education, 2012b). For females, 43% of White, 28% of Hispanic/Latino, and 18% of Black females scored in the top third compared to 23% White, 39% Hispanic, and 51% Black females scoring in the lowest third (Florida Department of Education, 2012b).

Other science district administration results

When the past academic performance for the North Florida school under investigation was examined during the 2010-2011school year, FCAT Science scores from the 137 11th graders, showed that the average mean scale score was 278, which is below the state average of 307 (Florida Department of Education, 2011a). While only 26% of students scored at proficiency or higher compared to the state average of 40% at the North Florida school (Florida Department of Education, 2011a), student data indicates an increase in student performance by 9% from the previous 2010 spring administration (Florida Department of Education, 2010).

This performance is not particularly surprising when the academic performance of the lower grades that feed into the North Florida high school is examined. According to Florida Department of Education (2012d), only 22% of students scored at proficiency or above compared to 46% of eighth grade students statewide during the 2012 academic year administration of the 8th grade science FCAT 2.0.Results from the fifth-

grade assessment also revealed that 26% of students scored at proficiency or above, compared to 51% statewide (Florida Department of Education, 2012c).

Thus the future students in the North Florida high school have underperformed in science in general when compared with the state and have also underperformed in the domains of the foundations of science and biological content (Florida Department of Education, 2012c, 2012d). Thus at the 5th and 8th grade level, students already exhibit a deficit in content area knowledge that serve as the background knowledge needed for more advanced understanding of biological content. What is concerning is that it is these students that will be tested upon their enrollment in biology I or an equivalent course at the secondary level, and their deficits in lower grades will serve as a detriment to their performance.

Statewide, student performance in science as measured by the FCAT Science in grades 5, 8, and 11 has improved in the state of Florida between 2003 and 2011 (Florida Department of Education, 2011). The percent of students scoring at proficiency and above levels at the fifth grade increased from 28% in 2003 to 51% in 2011, at the eighth grade level from 28% in 2003 to 46% in 2011, and from 33% in 2005 to 40% in 2011 for eleventh graders (Florida Department of Education, 2011b). Despite improvements on this indicator, another indicator of student achievement in science fails to illuminate such improvements. According to the National Center for Education Statistics (2012), the 2011 data on 8th grade science performance indicates that 72% of Florida students scored at or below basic. Furthermore, the average score of Floridian 8th graders did not change significantly between 2009 and 2011 and remained below the national average, resulting in Floridian students scoring below 37 other states

(National Center for Education Statistics, 2012). Therefore these indicators of student performance signify that nationally, and particularly in the state of Florida, that students are grossly unprepared for future science coursework or more advanced technological careers that will become increasingly commonplace in the future.

Computerized Testing

One form of test pollution that has been documented by empirical research which may skew the accuracy of Biology end-of-course assessments is the use of computerized testing. To address stakeholder concerns regarding computer-based assessments, Florida Department of Education (2006) published a paper on the empirical comparability of paper- and computer-administered assessments. According to Florida Department of Education (2006), test scores derived from computer-administered and paper-administered tests are comparable. Although the state agency cited research that indicates that some students perform better on a computer-administered test and others on its paper-based counterpart, the inconsistency was explained by highlighting the dissimilar computerized test administration systems (Florida Department of Education, 2006).

Comparability of computer-based testing

The state agency then further cited previous research studies, the majority (seven) of which indicate the comparability of test scores in biology attained via computer and paper-based administration, one citing computers as the more difficult administration mode (Cerillo & Davis, 2004), and three citing paper-based as more difficult (Russell, 1999; Russell & Haney, 1997, 2000). Pomplum, Ritchie, and Custer (2006) revealed that students eligible for free lunch performed better on paper-and-pencil tests. However, the discrepancy in scores between paper-and-pencil tests and

computerized tests decreased as student grade level increased. Pomplum, Ritchie, and Cruster (2006) explained that this reduction in the score discrepancy may be due to increased student familiarity with computers at each grade level. Despite this result, this study has limited applicability to the Biology 1 end-of-course assessment as the participants were enrolled in the lower grades at the elementary level where reading proficiency serves as a confounding variable due to the age of the participants.

Russell (1999) and Russell and Haney (2000) however, focused on computer-based administration of open-ended test questions; a format that is not utilized in the Biology end-of-course assessment; thus both studies are not applicable. On the other hand, Russell and Haney's 1997 study focused on the comparability of computer and paper-and-pencil test administration on multiple choice test items and revealed that the mode of administration is inconsequential. Mead and Drasgow's (1993) meta analysis of 29 studies also found no significant difference between paper-based and computer-based test administration of multiple-choice tests, particularly for timed power tests such as the Biology I end-of-course assessments.

More recent studies however have indicated the equivalence in student performance on online and paper-based tests (Escudier, Newton, Cox, Reynolds, & Odell, 2011; Frein, 2011; Horne, 2007; Kingston, 2009). Research has also indicated that some students have performed better in the online medium, completed more questions, and the majority of students expressed a preference for the online medium (Escudier, Newton, Cox, Reynolds, & Odell, 2011; Frein, 2011; Horne, 2007). According to Kingston (2009), the discrepancies in studies on the comparability of computer-based and pencil-and-paper tests may be partially due to the changes in the quality of

computer-based test administration systems as well as the computer experience of students.

For example, older computer systems required that students scroll up and down and from side to side to read an entire passage or large item which made the computer platform less user-friendly and more difficult. Newer systems however, have minimized scrolling of stimuli to a single dimension at most and the majority of newer systems avoid the scrolling of items (Kingston, 2009). According to Kingston (2009), reading from a page is different than reading while scrolling as while reading a page, spatial memory clues are used in which students remember seeing information regarding a particular question in a specific location on a page. Therefore, the student can quickly return to that spot to find the information. However, in a traditional computer-administered system, these parallel clues are unavailable as the spatial frame of reference is changed while scrolling. Thus some newer systems enable students to highlight text as they read or provide other cues that reduce score differences due to task differences (Kingston, 2009).

Other empirical evidence on computerized testing

Other factors that were examined empirically in the Florida Department of Education (2006) publication include: student preference for computer-based testing, inconclusive evidence regarding student computer experience and its effect on academic performance, no gender-based comparability differences, and text presentation and slow response times due to internet connection speed or school network constraints has a negative effect on comparability (Florida Department of Education, 2006). However, caution must be employed when examining student results from the empirical analysis that was used as the basis of the decision making process

employed by the Florida Department of Education (2006) because a paucity of empirical research exists specifically regarding biology multiple choice test items, some of which contain long passages, that are timed and non-adaptive. Most research regarding computer-based assessments fail to focus on the specifics regarding this type of assessment although the test is considered high stakes.

Computer platform for the Biology EOC

In fact, the Biology I end-of-course assessment provides the highlighting function for students in addition to other tools such as: a "reset" button that removes the selected answer, a "review" button that enables the student to flag a question for review at a later time, a "go to" button that allows the student to access the question marked for review on the item review screen, an "eliminate choice" tool that enables the student to cross out answer choices deemed incorrect, an eraser tool to remove highlighting or remove an X from an eliminated answer choice made with the eliminate choice tool, a calculator, and a help button (Pearson Education, 2013).

To further ensure that students are not negatively impacted from the computerized platform, students are also provided four-page, hard-copy work folders as scratch paper (Florida Department of Education, 2012i). Students also can utilize the practice test available online to gain familiarity with the software (Florida Department of Education, 2012i). For students at a disadvantage, paper-based testing is available for students with disabilities based on their Individual Educational Plans (IEPs) or Section 504 plans (Florida Department of Education, 2012i). Thus, concessions have been made by the Department of Education to ensure that scores are representative of student proficiency by: allowing students who need the alternative option are provided with paper-based testing; and that those that are not only familiar with the software, an

option to gain familiarity with it, while ensuring that the software utilized is user-friendly and contains tools that minimize the change in testing format.

The Importance of Science Education

> Second only to a weapon of mass destruction detonating in an American city, we can think of nothing more dangerous than a failure to manage properly science, technology, and education for the common good over the next quarter century (The United States Commission on National Security/21st Century, 2001a, p. 30).

The aforementioned quote from the United States Commission on National Security/21st Century underscores the Commission's view of education as critical to ensuring the nation's security and its future. That perspective has become increasingly prominent due to the September 11, 2001 attacks in which the strategic role of science and technology in the post-Cold War era has changed (National Science Board, 2004). Of concern has been the role of foreign students, scientists, and engineers within the United States science and engineering system; determining the appropriate degree of balance between security and openness within scientific communication; the degree of contributions that research and development are capable of making in domestic security; and the course of federal research and development initiatives (National Science Board, 2004).

Thus, in order to sufficiently meet the scale and nature of scientific and technological advances in society, the need for human capital is critical (The United States Commission on National Security/21st Century, 2001b; Trefil, 2008). Furthermore, the ability of Americans to compete globally is highly dependent upon their knowledge of science and math and their ability to apply this knowledge to emerging technologies (U.S. Department of Education, 1989). Thus, to be competent, citizens and workers require a sound understanding of science and mathematics (National Science

Board, 2004; Trefil, 2007; 2008). Science and mathematics are considered essential to the decisions made regarding jobs, use of resources, health, and everyday activities of consumers (U.S. Department of Education, 1989). Furthermore, the products, services, standard of living, and economic and military security needed to sustain Americans at home and globally will be derived from mathematics and the sciences (National Commission on Mathematics and Science Teaching for the 21st Century, 2000). Consequently, in an integrated and global economy, math and science will serve as the core forms of knowledge that are required by innovators, producers, and workers to solve unforeseen problems and determine America's future (National Commission on Mathematics and Science Teaching for the 21st Century, 2000; Trefil, 2007).

Reasons Cited for Shortage in Science Personnel

However, despite this need for the highest quality human capital in science, mathematics, and engineering by the nation, it is currently not being met by the educational system (The United States Commission on National Security/21st Century, 2001b). This system operates under the notion that the elementary and secondary schools are ensuring that students acquire knowledge in science and mathematics (National Science Board, 2004). One reason cited as contributing to this problem of an unprepared citizenry is the view that the American K-12 educational system fails to prepare students as it should for college or the commercial sector system (The United States Commission on National Security/21st Century, 2001b) due to the minimization of science and mathematics instruction, particularly in the early grades, and the lack of qualified personnel in these content areas in teaching positions (U.S. Department of Education, 1989).

Additionally, the lack of importance placed on education, according to Christensen, Horn, and Johnson (2011), can be partially understood when one understands the effect of prosperity on the nation. The authors explained that the educational malaise that is affecting this nation is in part, due to the country's stability and prosperity (Christensen, Horn, & Johnson, 2011). For students residing in a developing country that is becoming an increasingly industrial-based economy, studying subjects such as science, math, and engineering offer rewards in which the individual can escape from poverty. The freedom that results from living in a prosperous nation results in students having greater freedom to study subjects they find fun and intrinsically motivating. This proposed explanation appears fitting based on results of surveys that indicate that positive student attitudes toward science (Gibson and Chase, 2002; U.S. Department of Education, 1989) and mathematics decline as students advance in coursework (U.S. Department of Education, 1989). Additionally, most parents and students subscribe to the notion regarding the unimportance of science and mathematics for the majority of students and that high achievement in these content areas are due to factors other than effort and hard work (U.S. Department of Education, 1989).

Catalyst for the Improvement of Science Education

This may, in part, explain why the percent of students scoring at proficiency levels in science decreases from the 4rth to the 12th grade (National Science Foundation, 2011) and why foreign students account for 57% of all doctorates in engineering, 54% in computer science, and 51% in physics (National Science Foundation, 2012a). Moreover, of the 5 million degrees earned in science and engineering worldwide, 23% of those were obtained by students in China, 19% from

those in the European Union, and only 10% within the United States (National Science Foundation, 2012a). China, in fact, has awarded the most number of doctoral degrees in natural sciences and engineering since 2007, thus overtaking the United States as the world leader (National Science Foundation, 2012a).

Consequently the dissatisfaction of parents, policymakers, legislatures, and educators regarding mathematics and science underachievement, has resulted in numerous efforts to both reform and improve schools (National Science Board, 2004). National effort had been galvanized since the appearance of Sputnik to improve mathematics and science education (National Commission on Mathematics and Science Teaching for the 21st Century, 2000). The publication of *A Nation At Risk* urged greater accountability of schools, as well as higher academic standards, and better teacher preparation as a means of improving schools (National Commission on Excellence in Education, 1983). Other reports and commissions thereafter established lofty goals, such as stating that by the year 2000, U.S. students would rank "first in the world in science and mathematics achievement" by strengthening math and science education in early grades, increasing the number of math and science teachers with substantive backgrounds in those fields by 50%, and significantly increasing the number of U.S. undergraduate and graduate students with degrees in math, science, and engineering; especially women and minorities (U.S. Department of Education, 1989). In response to this announcement, the National Commission on Mathematics and Science Teaching for the 21st Century announced in 2000 that U.S. students were "devastatingly far from this goal by the time they finish high school" (p. 10).

According to American Association for the Advancement of Science (AAAS) Project 2061 (1990), two disparate and growing public concerns have motivated the majority of education reports in the 1980s. One such concern is the seeming decline of America's economy. When the domestic affluence and international power, as evident in scientific and technological domination of America is compared with other countries, particularly Japan; they appear to be weakening. The other concern is based on trends in U.S. public education that include low test scores, the avoidance of mathematics and science by students, a weakened teaching staff characterized by demoralization, a comparatively low learning expectation with other technologically advanced nations, and the ranking of U.S. students near the bottom in mathematics and science on international studies. Both the reports and coverage by the mass media of such reports have highlighted the deficiencies in education, and thus the nation became aware of the crisis in American education. However, the most powerful argument to improve science education is its role in "liberating the human intellect" although much of the public discussion has been focused on more tangible, utilitarian, and immediate explanations (American Association for the Advancement of Science (AAAS) Project 2061, 1990).

School reform thereafter became increasingly strengthened under the Federal No Child Left Behind (NCLB) Act of 2011 in which school accountability measures were implemented by requiring that schools demonstrate progress in student achievement in core subject areas using high-stakes testing to measure learning. The act specified to each state, the immediate development of standards in math and science by the academic year 2005. Furthermore, the NCLB Act mandates students in grades 3 through 8 to be assessed every year beginning in the 2005 academic year for math and

2007 for science. Sanctions will be imposed on schools that fail to show improvement in student achievement for all students after they initially received assistance (No Child Left Behind Act of 2001, 2002).

Despite such attempts, U.S. students continued to perform poorly in both science and mathematics. In his attempt to address the underachievement of American students in math and science, President Obama instituted the first of two White House Science Fairs as a means of fulfilling his commitment to improve American student performance from the middle to the top of the pack over the next decade (The White House, Office of the Press Secretary, 2010). The President also indicated the importance of Science, Technology, Engineering, and Math (STEM) education in providing the new foundation that will determine America's future prosperity (The White House, Office of the Press Secretary, 2010). In order to achieve this goal, his administration allocated $ 4 billion to Race to the Top (RTT) initiatives in which states were empowered to develop a comprehensive strategy to improve student achievement in STEM subjects, especially for women and underrepresented minorities (The White House, Office of the Press Secretary, 2010).

This decision by the Obama administration to allocate fiscal resources to address the issues of student underachievement in science represents one of many attempts by the federal government to improve the science performance of American students (U.S. Department of Education, 2004). This underperformance of American students was significant enough to warrant a statement within the 2011 State of the Union address (The White House, Office of the Press Secretary, 2011) and is viewed as limiting the nation's ability required for global economic leadership as well as homeland security in

the 21st century (U.S. Department of Education, 2004). Thus, America's children are viewed as failing to adequately respond to the challenges for the 21st century but its potential as well (National Commission on Mathematics and Science Teaching for the 21st Century, 2000).

Reading and Science Achievement

Of the many challenges within the education sector, the proficiency level of U.S. students in reading is a disconcerting one. Considering that reading is a requirement on assessments, one can reasonably assume that reading proficiency can impact test scores. What remains unclear is whether test scores are affected by student inability to understand the question being asked or in the case of an incorrect response, whether the reason is due to lack of knowledge or an inability to comprehend the test item (Homan, Hewitt, & Linder, 1994).

Importance of Reading

Of the many indicators of reading proficiency at various grade levels, U.S. students have consistently engaged in paltry performance (ACT, 2010; College Board, 2011; Cromley, 2009; Fleischman, Hopstock, Pelczar, & Shelley, 2010; Florida Department of Education, 2012e; National Center for Education Statistics, 2011; Torgesen, 2006). Considering such dismal student performance on fundamental literacies, it is feasible to state that the majority of high school students and graduates are lacking the reading skills needed to successfully traverse the Information Age as well as navigate content area learning. Furthermore, it is argued that basic skills are no longer sufficient to successfully navigate the post secondary world (Meltzer, Cook Smith, &Clark, 2002) and consequently multiple indicators of student performance such as the criticisms of employers, standardized test scores, standards-based assessment

results, and the 2011 NAEP (National Center for Education Statistics, 2011) paint a melancholy picture on the competencies of our students.

Strong literacy skills are necessary for adolescents to navigate the world, communicate effectively, understand academic content, and partake in cultural communities (Meltzer, 2001). This concept was held by one of the greatest lyricist of American Democracy, Thomas Jefferson, who noted the direct relationship between literacy, citizenship, and successful self-government. Good citizenship, in his opinion, was not capable without the knowledge and discernment that accompany literacy. In addition to participating in society, literacy enables the individual to access opportunities for economic vitality, and individual or personal fulfillment. Although various cultural and language minority groups have rich written and verbal literacy indicated in research, the lack of literacy in the larger society fails to create equal opportunities of access to resources in that society and consequently to the privileges, rights, and responsibilities of American citizenship in a democratic society (Meltzer, Cook Smith, & Clark, 2002).

Literacy Rates of American Students

An examination of various assessments reveals that in general, literacy levels of all subsets of the school population is low, and particularly among certain subsets. For example, one test, the ACT, defines "college readiness" as the probability of students earning a grade of "C" or higher at 75% or a 50% chance of earning a B or higher in first-year college courses such as College Algebra; English Composition; History, Psychology, Sociology, Political Science or Economics; and Biology (ACT, 2010). Of the 2010 ACT-tested high school graduates, only 24% met all four ACT College Readiness Benchmarks- thus only 24% of students were academically prepared for first-year college courses such as College Algebra, English Composition, social sciences, and

Biology (ACT, 2010). Of the students who completed the 2010 examination, only 62% of White, 61% of Asian, 34% of Hispanic, and 21% of African American high school graduates met College Readiness Benchmarks in Reading. Thus, the access to economic vitality obtained through obtaining a undergraduate degree is differentially available to students due to academic performance in reading and its effect on subsequent collegiate course grades.

Other academic indicators paint a similar picture of general paltry performance in reading and differential performance based on ethnicity, gender, and socioeconomic status. The SAT, which is a predictor of student performance in college; indicates that the average reading score has dropped since the 1986-1987 administration (U.S. Department of Education, National Center for Education Statistics, 2012). Since the 1986 administration, White students consistently perform higher than the average score while their Black and Hispanic counterparts score lower than the average (U.S. Department of Education, National Center for Education Statistics, 2012). Specifically, the average score of Black (428) and Hispanic (451) students was lower than the national average (497) on SAT critical reading although their White (528) and Asian (517) peers scored higher (College Board, 2011). Furthermore, as family incomes increased, the average score on the critical reading section also increased (College Board, 2011).

The poor academic performance of high school students however, is not unforeseen when one considers the performance of students at the elementary level. According to the National Center for Education Statistics (2011), since 2007, 67% of students were scoring at the basic level. When eighth graders were assessed, the

percent of students scoring at basic levels increased to 76% (National Center for Education Statistics, 2011). Therefore since 2007, student performance at the fourth-grade level has remained relatively unchanged, thus the majority of the students continue to lack proficiency in reading. At the eighth-grade level, even greater numbers of students are failing to demonstrate proficiency.

Trends in Student Literacy

Despite this alarming trend, students from high- and lower-income families increased academic performance (National Center for Education Statistics, 2011), thus there is some improvements on academic gauges worth celebrating. When disaggregated, certain trends become readily apparent. Students who read more often score higher than their peers who read less often (National Center for Education Statistics, 2011). Additionally, students who scored above the 75th percentile were predominantly White (71%) and read for fun almost every day (60%). For students who scored below the 25th percentile, 38% read for fun almost every day and 74% were eligible for free/reduced lunch (National Center for Education Statistics, 2011).

Another readily discernible trend is the difference in achievement obtained by students of different ethnicities. For example, the mean score of White fourth-grade students (231) was higher than Black (205) and Hispanic (206) students in 2011 (National Center for Education Statistics, 2011). When the performance of fourth-grade Black students was further examined, 84% of Black students scored below or at the basic level. When compared with their White counterparts, only 57% scored below or at the basic level (National Center for Education Statistics, 2011). Thus Black fourth-grade students accounted for proportionately higher levels of students scoring at or below basic levels. They also accounted for proportionately lower levels of students scoring at

advanced levels when compared with their White counterparts. At the eighth-grade level, similar results in performance were also revealed (National Center for Education Statistics, 2011). Other discernible trends regarding socioeconomic status were also revealed. When socioeconomic status data was also analyzed at both grade levels, students eligible for free lunch scored significantly lower than those eligible for reduced-price lunch and lower than those not eligible (National Center for Education Statistics, 2011).

International comparison of student literacy

When viewed internationally, U.S. 15 year-old students obtained an average score of 500 on the combined reading literacy scale, higher than the 493 OECD average (Fleischman, Hopstock, Pelczar, & Shelley, 2010). Although this may indicate that students are performing at relatively proficient levels, when compared with other countries, 15 OECD and non-OECD countries earned higher average scores. Furthermore, only 30% of U.S. students scored at or above proficiency in locating and organizing "several pieces of embedded information", "interpreting the meaning of nuances in language", "understanding and applying categories in an unfamiliar context", and demonstrating "an accurate understanding of long or complex tests whose content or form may be unfamiliar" (Fleischman, Hopstock, Pelczar, & Shelley, 2010, p. 10).

Trends in student literacy in Florida

Specifically in the state of Florida, the State Board of Education established new Achievement Level standards for FCAT 2.0 Reading on December 19, 2011 (Florida Department of Education, 2012e). Florida Comprehensive Assessment (FCAT) 2.0 is considered to be more demanding in its content standards and more rigorous in its achievement standards than the former FCAT version of the Reading test (Florida

Department of Education, 2012e). Consequently, the scores reported are different from previous scales for the 10th grade assessment. In response to the change, new Achievement Levels were developed as follows for all FCAT 2.0 and EOC assessments:

- Level 5: Students at this level demonstrate mastery of the most challenging content of the Next Generation Sunshine State Standards.
- Level 4: Students at this level demonstrate an above satisfactory level of success with the challenging content of the Next Generation Sunshine State Standards.
- Level 3: Students at this level demonstrate a satisfactory level of success with the challenging content of the Next Generation Sunshine State Standards.
- Level 2: Students at this level demonstrate a below satisfactory level of success with the challenging content of the Next Generation Sunshine State Standards.
- Level 1: Students at this level demonstrate an inadequate level of success with the challenging content of the Next Generation Sunshine State Standards. (Florida Department of Education - Office of Assessment, 2012j).

In its first year of implementation, only 51% of 10th grade students scored at proficiency or above on the FCAT Reading 2.0 assessment (which is considered level 3 or higher) (Florida Department of Education, 2012e). Specifically, 20% scored at level 1 (inadequate level of success), 30% at level 2 (below satisfactory), 22% at level 3 (satisfactory level), 19% at level 4 (above satisfactory), and 10% at level 5 (mastery of most challenging content) (Florida Department of Education, 2012e).

Therefore, one can derive a general consensus that on scales of a statewide, national, and international level, American students are failing in disproportionate numbers to achieve basic literacy skills regardless of years spent in formal education. This failure to successfully navigate basic literacy skills has dire consequences on content area learning. Although content areas vary in their literacy demands (Grossman & Stodolsky, 1995), the areas of knowledge and skills necessary to advance literacy

between grade 4 and 12 include: a) automaticity of word recognition, b) the dramatic expansion of vocabulary, c) the growth of conceptual knowledge and understanding, d) the increase of thinking and reasoning skills, e) the development of self-regulated reading comprehension strategies, and f) the acquisition or maintenance of a motivation and interest in broad and deep reading (Torgesen, 2006).

Results of Research on Reading Proficiency and Science Achievement

Currently, a limited number of studies exist regarding reading proficiency and its subsequent effect on science proficiency, yet assessments and activities in science generally require that students read text (Norris & Phillips, 2003; Yore, Bisanz, & Hand, 2003). Furthermore, students may fail assessments due to their inability to read or understand the test, and not because of lack of knowledge with the content (Roe, Stoodt, & Burns, 1991). Studies on reading and science achievement generally fall into one of three categories: a few correlational studies, research on language-based science instruction, and research on the comprehension of science text (Cromley, 2009). However, despite the available research on the impact of reading on science achievement, only one study to date has specifically investigated the impact of reading proficiency on biology achievement. Thus researchers are relying primarily on results obtained from comprehensive science assessments to make plausible inferences regarding the influence of reading proficiency on biology achievement.

The only study that focused primarily on biology content by Haught and Wall (2004) of 730 medical school students revealed an interesting finding. The methodology involved participants completing the Nelson-Denny Reading Test that assesses reading vocabulary, reading comprehension, and reading rate. Prior scores on the Medical College Admissions Test (MCAT) that included a verbal reasoning and biological

science section were collected. Results from the study reported that the Nelson-Denny Reading Test and its 3 sections were significant predictors of MCAT verbal reasoning. The vocabulary portion of the Nelson-Denny Reading Test was positively correlated with the MCAT biological science score ($r = 0.27$, $p < .01$), the MCAT verbal reasoning score, ($r = 0.53$, $p < .01$), and the United States Medical Licensing Examination (USMLE) score ($r = 0.14$, $p < .01$).

An increase in the vocabulary score on the Nelson-Denny Reading Test was commensurate with an increase in the MCAT and USMLE exam scores. The comprehension portion of the Nelson-Denny Reading Test was positively correlated with the MCAT biological science score ($r = 0.12$, $p < .01$) and the MCAT verbal reasoning score, ($r = 0.41$, $p < .01$). The total reading portion of the Nelson-Denny Reading Test was also positively correlated with the MCAT biological science score ($r = 0.23$, $p < .01$) and the MCAT verbal reasoning score, ($r = 0.56$, $p < .01$). Additionally, reading vocabulary was a significant predictor variable of MCAT biological science scores. Therefore, vocabulary, among all other variables in the study, served as the most significant predictor of scores on other assessments such as MCAT and USMLE as well as proficiency in the biological sciences. This finding, while consistent with other reports using comprehensive science assessments, were not replicated in most studies of reading proficiency and general science achievement.

A cursory examination of anecdotal evidence of reading and biology assessments reveals a pattern of reading scores being relatively similar to biology scores. For example, on the SAT Biology subject test, out of a possible 800, the average score was 604 for the Ecology focus and 635 for the Molecular focus (College

Board, 2011). Of students who took those subject tests, the average score on the critical reading section was 596 and 609, respectively out of a possible 800 (College Board, 2011). Despite the relative similarity in scores from an observational account, a limited number of studies have used statistical methods in a more objective manner.

Visone (2009) investigated the relationship between reading and science achievement on a standardized test. Results indicated a moderate-to-strong positive relationship between the variables, with a range from 0.41 to 0.74. Another study by Visone (2010) sought to determine the reading issues associated with a standardized test in science by analyzing the self reports of 11th grade students. Qualitative results revealed that some of the science questions contained too much information that served to distract the students from their purpose, the use of sophisticated vocabulary, failure to understand scientific terminology or concepts, lack of background knowledge (Visone, 2010).

Maerten-Rivera, Myers, Lee, and Penfield (2010) revealed that reading and mathematics accounted for 58% of the variation in science achievement among student-level factors. According to the researchers, for every one-unit increase in reading achievement, a resulting average increase of 0.51 points occurred in science performance. Additionally, reading performance accounted for 25% of the variation in science achievement (Maerten-Rivera, Myers, Lee, & Penfield, 2010). Romance and Vitale (2008) reported an increase in student achievement on both the Iowa Test of Basic Skills comprehension subtest and the MAT science achievement test by increasing science content knowledge through reading, writing, inquiry, and discussion without any explicit instruction in vocabulary or comprehension strategies.

Other increases in science achievement have also been identified due to the implementation of broad reading interventions in which students are instructed in reading strategies, vocabulary, background knowledge, and writing (Fang & Wei, 2010; Guthrie, Wigfield, Barbosa, Perencevich, Taboada, Davis, Scafiddi, &Tonks, 2004; Reaves, 2000; Shymansky, Yore, & Anderson, 2004; Vitale & Romance, 2012; Yore, Bisanz, & Hand, 2003). When the results of these two studies are analyzed, it is therefore not startling that Dempster and Reddy's (2007) study revealed that science performance is relatively dependent on reading skills.

One study, conducted by O'Reilly and McNamara (2007) examined how well science knowledge, reading skill, and reading strategy knowledge predicted science achievement. Using data from 1,433 students, they revealed that the correlation of science knowledge and reading skill was 0.577. Additionally, science knowledge, reading skill, and reading strategy knowledge accounted for 36% of the variance in answers to multiple-choice questions, with reading skill and science knowledge contributing significantly to the model. The model of science knowledge, reading skill, and reading strategy knowledge also contributed to 46% of the variance in which reading skill and science knowledge served as significant predictors in the model. Reading strategy knowledge however was not significant.

Thus, reading skill and science knowledge served as significant predictors of science achievement (O'Reilly & McNamara, 2007). Reading skill also significantly improved science proficiency for students with low levels of science knowledge so that they scored as high or higher than their less skilled, higher knowledge readers; thus compensating for low knowledge (O'Reilly & McNamara, 2007). For higher knowledge

students, a high level of reading skill significantly improved science achievement (O'Reilly & McNamara, 2007).

Another study conducted by Medina and Mishra (1994) examined the relationship between the reading achievement and academic performance in science of 518 Mexican American students residing in Arizona. Although the tests were conducted in the native language of the students, results nevertheless indicated a significant positive relationship between native reading and native science performance (0.61, $p < 0.001$) (Medina & Mishra, 1994). Three other studies of English-speaking high school students also similarly reported high correlations. Nolen (2003) reported a strong correlate between scores on the district science achievement test and reading comprehension of $r = 0.60$. Demps and Onwuegbuzie (2001) revealed a correlation of 0.80 between the scores on the Iowa Test of Basic Skills (ITBS) reading assessment and the high school science graduation test. Additionally, a correlation of 0.78 between the high school science graduation test and the high school language arts graduation test was also found (Demps & Onwuegbuzie, 2001).

Last, another study sought to analyze the relationship between reading comprehension and science proficiency in more than 40 countries using 15-year-old-students who completed the PISA reading and science literacy assessments in 2000, 2003, and 2006 (Cromley, 2009). Using both scores obtained from each year of the reading and scientific literacy assessments, Cromley (2009) determined a mean correlation of 0.840 between reading and science and a range of 0.675 to 0.916 at a significance level of 0.001 across 43 countries using PISA 2000 data. This significant variation among the 43 countries was as a result of correlations that were the highest in

countries that were in the top one half on mean reading achievement, such as The Netherlands, the United States, and the United Kingdom; and low among the six countries that were the lowest achieving in reading such as Indonesia and Peru, except Latvia. Therefore, a stronger relationship between reading scores and science scores occurred in countries with higher mean reading performance than countries with lower mean reading performance.

When Cromley (2009) replicated the study using 2003 data from 276,192 students from 41 countries, a correlation of 0.805 was determined as well as a range of 0.599 to 0.892 at a 0.001 significance level. Data obtained from the 2006 administration of 389,750 15-year-olds also revealed a correlation of 0.819 at a 0.001 significance level (Cromley, 2009) despite the test designers deliberately requiring less reading (OECD, 2007). Cromley (2009) also found a range of 0.603 to 0.902 at a 0.001 significance level among 56 countries. Similar to previous findings in the study, high correlations were present from countries in the top one half on mean reading achievement and low correlations among the lowest achieving. Thus, reading and science scores were stronger in countries with higher mean reading performance (Cromley, 2009). Other research using PISA data indicated a 0.889 correlation between latent reading factors and latent science factors for retrieving information, 0.890 for interpreting texts, and for reflection and evaluation, a 0.840 correlation (OECD, 2002).Thus of the existing literature on the relationship between reading proficiency and science proficiency, all studies at the secondary level indicate a strong positive relationship.

Cromley (2009) concluded that the significant correlation between reading comprehension and science proficiency is due to good reading comprehension causing scientific proficiency as good reading comprehension should result in a good understanding of scientific texts and tests. Other research findings concur with this conclusion as acquiring scientific information through knowledge and skills is undertaken through reading (Ediger, 2005). Cromley (2009) also proposed that increases in science proficiency could result from the products of extensive reading experience such as background knowledge, reading comprehension strategies, inference, and general vocabulary. Other research has also indicated that because proficient readers engage in more reading (Bray, Pascarella, & Pierson, 2004; Torgesen, 2006) and read more widely (Smith, 2000; Torgesen, 2006; Wingfield & Guthrie, 1997), their exposure to scientific text would be more than their less proficient peers.

Hence, a myriad of factors contiguous with reading contribute to general scientific proficiency. For purposes of this review, I will focus solely on the two reading requirements that limited studies in biology have demonstrated are responsible for directly improving biology achievement: background knowledge and vocabulary.

Background knowledge

Background knowledge serves as an important impetus of increasing reading literacy. Marzano (2004) indicated that the ability to learn new information is dependent upon what students already know. Thus, background knowledge is necessary to activate existing knowledge needed integrate new material into preexisting information (Feathers, 2004). Utilizing background experiences and knowledge occurred more often in the better readers who used the text in conjunction with their experiences and

knowledge to reformulate old ideas. Poor readers were less likely to use the text to rethink their experiences (Langer, 2001). Based on these findings, it is therefore not surprising that students from minority and low socioeconomic backgrounds typically were characterized as poor readers as they lacked ample background knowledge, and thus earned paltry scores on reading assessments (Fisher, Ross, & Grant, 2010; Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Reeves, 2000).

Consequently, lack of use of this salient element in reading is attributable to deficiencies in reading comprehension (Best, Rowe, Ozuru, & McNamara, 2005; Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Guthrie, Wigfield, Barbosa, Perencevich, Taboada, Davis, Scafiddi, &Tonks, 2004; Ozuru, Dempsey, & McNamara, 2009). In fact, research has shown that the better predictor of science text comprehension is prior knowledge, not reading skill (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Ozuru, Dempsey, & McNamara, 2009). Prior knowledge about biology was positively correlated with overall comprehension of science text and accounted for a large portion of variance in performance on comprehension questions above and beyond reading skills (Ozuru, Dempsey, & McNamara, 2009). The effect of prior knowledge is also larger on questions that require more extensive integration of information while the effect of reading skill was larger on questions that were text-based (Ozuru, Dempsey, & McNamara, 2009).

Thus scientific learning and performance is hampered by limited background knowledge. Learning has also been cited as a challenge due to the nature of scientific texts. Research has indicated that students have difficulties understanding expository texts and in particular those scientific in nature (Best, Rowe, Ozuru, & McNamara, 2005;

Ozuru, Dempsey, & McNamara, 2009). Some of the proposed explanations for this difficulty is due to student difficulties in generating inferences required to understand science as they possess a low level of prior knowledge and inadequate reading strategies (Best, Rowe, Ozuru, & McNamara, 2005; Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Ozuru, Dempsey, & McNamara, 2009). Consequently, readers are often deficient in a sufficiently developed mental model that is representative of the overall conceptual relationship between relevant concepts such as hormones and tropism (Kendeou & Van Den Broek, 2007; Ozuru, Dempsey, & McNamara, 2009).

Another reason cited for the difficulty in understanding scientific text is due to the manner in which experts on the topic convey information to less knowledgeable readers (Best, Rowe, Ozuru, & McNamara, 2005; Norris & Phillips, 2003). The process of creating a textbook involves one or a few expert teachers who write the text and subject-matter experts and veteran teachers who edit and review it. Hence, the textbook by design “reflects the architecture of knowledge in its domain or field” (Christensen, Horn, & Johnson, 2011, p. 128). Thus “instructional materials in the school curriculum are developed by and taught to a ‘dominant intelligence’- the type of brain whose wiring is most consistent with the methods used to solve problems in the field, as the domain experts have framed them” (Christensen, Horn, & Johnson, 2011, p. 128-129).

Furthermore, the writing style in science texts is characterized by "low-cohesion" in which students must generate many inferences and fill in conceptual gaps (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Ozuru, Dempsey, & McNamara, 2009). According to Best, Rowe, Ozuru, & McNamara (2005), many science textbooks often omit information that the authors assume are prior knowledge for their readers; thus

exacerbating previously existing reading deficiencies due to poor background knowledge. Since students lack specific knowledge on concepts such as osmosis, and gravity; their ability to generate inferences that links concepts across and within sentences is limited (Best, Rowe, Ozuru, & McNamara, 2005; Kendeou & Van Den Broek, 2007; Ozuru, Dempsey, & McNamara, 2009). Thus students experience a fragmented and isolated understanding of the text which prevents students from obtaining a coherent mental representation of the overall information (Best, Rowe, Ozuru, & McNamara, 2005; Kendeou & Van Den Broek, 2007; Ozuru, Dempsey, & McNamara, 2009). However, this omission of background knowledge by authors has been shown to be beneficial with reading comprehension with highly-skilled readers and detrimental to their less skilled peers (Ozuru, Dempsey, & McNamara, 2009).

Research has in fact indicated that the most common problem for older readers is comprehension of text because some students may not possess sufficient fluency to aid in comprehension, others are lacking the strategies to comprehend the text, while others are not versed in employing the use of these strategies in a variety of situations (Biancarosa & Snow, 2004). Thus, sophisticated reading comprehension skills are necessary at the secondary school level to successfully navigate the demands of more challenging academic expectations (Meltzer, Cook Smith, & Clark, 2002). Linguistic features in academic subject texts, particularly science, contribute to the abstract and dense nature of these texts that are unfamiliar to the texts used by students at earlier ages. These features contribute to comprehension challenges, especially for struggling readers and English-language learners; and may be overlooked by proficient adults (Fang, 2006; Ozuru, Dempsey, & McNamara, 2009).

These challenges are evident in the research conducted by Torgesen (2006) that investigated passage length according to grade level on the FCAT. Results indicated that the ability of students to understand longer and more complex ideas and sentences as grade levels increased was poor (Torgesen, 2006). Using 2005 FCAT data, Torgesen (2006) reported that as Floridian students increased in grade level, the proficiency levels in reading decreased. This is partially due to the increased skills and knowledge required to meet each grade level standard, such as more extensive use of background knowledge, being able to handle lengthier sentences with complex ideas, increased vocabulary demands, and the automaticity required in recognizing new vocabulary (Torgesen, 2006).

Thus the comprehension process consists of two salient features in which background knowledge is activated and the conveyance of ideas occurs through printed words. Therefore limited background knowledge and good word-level processes (such as quick and effortless word identification and the meanings of keywords) if absent, serve as impediments in the comprehension process (Feathers, 2004; Neufeld, 2005; Marzano, 2004; Marzano & Pickering, 2005; Ozuru, Dempsey, & McNamara, 2009) and has been found to be an indicator of how well students learn content information (Marzano, 2004). Therefore, the following are erroneous independent assumptions- that poor readers are incapable of deriving meaning from text, or that they are unaware of utilizing a variety of sources to obtain knowledge, or that poor vocabulary is the basis of the differences in comprehension between low and high performing students. Instead, better readers have the knowledge of how to build on the information that they have grasped or understood (Langer, 2001).

Vocabulary

Vocabulary is also important in improving proficiency in reading as it involves the study of multiple relationships or the connection of ideas. Thus, previous vocabulary serves as a framework for future vocabulary (Smith, 1990). The acquisition of some vocabulary may occur more readily than others due to the frequency of its use in everyday scenarios and rarely require explicit instruction. These words constitute Tier 1 words which are comprised of words that commonly appear in spoken language. Tier 2 words consists of more sophisticated vocabulary of written text (academic words) that are used with regularity with mature language users but are encountered less frequently with students. Vocabulary that is limited to specific domains and appear only in isolated situations comprise Tier 3 words. As a result of the limited exposure of these words, students have difficulties internalizing the meanings of these words that are present in content area texts (Harmon, Hedrick, & Wood, 2005).

Examples of these words include medical, legal, and biology terms that are central to building knowledge and conceptual understanding within various academic domains although they are rarely used in the general vocabulary (Beck, McCaslin, & McKeown, 1980). To enhance vocabulary acquisition for Tier 3 words, instruction should encompass effective practices for general vocabulary as well as the unique features of the language of various content areas (Harmon, Hedrick, & Wood, 2005). It is therefore recommended in science that students should learn the specialized vocabulary of science when they are used in a unit of study (Fang, 2006). In fact, students need to learn to be attentive to the salient features inherent in the language of school science by creating an awareness of the similarities and differences between the

language of school science and the language of everyday life as well as understand how the language of science contributes to scientific thinking and thought (Fang, 2006).

Ethnicity and Science Achievement

According to measures of science achievement, students of minority backgrounds, on average; score lower than their White and Asian/ Pacific Islander counterparts and are disproportionately represented in scores of lower achievement (ACT, 2010; Bankston & Caldas, 1998; Florida Department of Education, 2012b; Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008; Maerten-Rivera, Myers, Lee, & Penfield, 2010; National Center for Education Statistics, 2012; Rathbun & West, 2004; Riegle-Crumb, Moore, & Ramos-Wada, 2010; Strand, 2012). For example, for ACT-tested high school graduates in 2010, only 44% of Asian, 36% of White, 14% of Hispanic, and 6% of Black students met ACT College Readiness Benchmarks in Science. Additionally, for U.S. fourth and eighth graders, average science scores of Black students were below the U.S. national average in 2007 and were the lowest of all ethnic groups while the average scores of their White and Asian counterparts were higher than the U.S. national average (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008). Hispanic students also scored below the U.S. national average at both the fourth and eighth grade level (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008).

Proposed Explanations for Disparities in Achievement

Based on these findings and other similar reports of underachievement in minority populations on every academic measure, some have wondered whether such differences are cultural or genetic in nature. The notion that underachievement is inherent within minority populations, especially within the Black population, and that

these populations are incapable of achieving at high levels due to their ethnic background is grossly incorrect. Instead, both cultural factors and systemic issues are being blamed for the achievement gap. According to Butchart (2010), despite the implementation of Jim Crow policies that mandated racial segregation, the legacy of slavery, the denial of resources- particularly social, intellectual, and financial; incongruent pedagogy and curriculum, cultural values antithetical to the Victorian values of school, and the ubiquitous segregated schools; African-Americans performed exceptionally well in formal education during the 1860s and 1870s.

Instead, the demise of African-American educational progress was due to the open White terrorism that aimed to destroy "the black dream of intellectual emancipation through education" (Butchart, 2010, p. 37) through resistance to black educational freedom via education-related violence such as harassment, ostracism, denial of physical resources to black schools, burning of schools or churches that housed schools, targeting the living quarters of teachers with violence, terrorism of black students, White indifference to the atrocities, and the use of violence, sometimes deadly; on teachers.

Easton-Brooks and Davis (2007) further corroborates this thesis by indicating that the economic and political history of African Americans is the root of the low educational performance of this subset of the population which schools cannot repair. Furthermore, socioeconomic accounted for significant variance in educational outcomes of African-American students but wealth accounted for greater variance in educational outcomes (Easton-Brooks & Davis, 2007). Thus, since more African-Americans occupy low socioeconomic and posses less wealth than their White counterparts (Caldas &

Bankston III, 1997; Snyder, Dillow, & Hoffman, 2008), the underachievement by African Americans is grounded in poverty rather than an outcome of race. In fact, research has indicated the interconnected nature of socioeconomic status and race (Aud, Fox, & KewalRamani, 2010; Caldas & Bankston III, 1997; House & Williams, 2000).

Furthermore, within the Black community, achievement varies. Pinder (2012) discovered that Afro-Caribbean students statistically scored significantly higher, on average, than their African American peers in science. The study also revealed that on average, students from Afro-Caribbean households received more assistance with homework, lived with their father, and had more books at home (Pinder, 2012). The study further revealed that the strongest predictors of science achievement were similar in both groups as reflected in the statements "parents discuss school progress" and "time spent in extra science lessons." However, the Afro-Caribbean group also cited, "performance motivation to do science" whilst the African-Americans additionally cited, "time studying science", and "science attitudes" (Pinder, 2012). Thus the study emphasized the importance of parental involvement in determining academic achievement in science.

This finding regarding parental involvement is consistent with Lee, Bryk and Smith's 1993 study that found that parental expectations for their children's achievement and the importance that parents place on education are positively and strongly related to academic performance. Even in high school, parents' beliefs, goals, and values concerning education and achievement strongly influence adolescents, who tend to incorporate these standards into their own. Of course, such influence may also be

negative, when parents do not value education or do not regularly enforce standards (Lee, Bryk & Smith, 1993).

Another interpretation of the result of Pinder's 2012 and Ogbu and Simons's 1998 studies states that the difference in immigration patterns (voluntary or involuntary) is responsible for part of the difference in educational outcomes. According to Ogbu and Simons (1998), voluntary (immigrant) minorities willingly immigrated to the U.S. unlike their African American counterparts who view their permanent presence as forced upon them by White Americans or the U.S. government during slavery. Furthermore, voluntary immigrants have a more positive view of American schooling as a means of gaining access to the greater opportunities available within the United States than their frame of reference which is "back home." Involuntary immigrants however hold a more negative view of White-controlled American schooling and a mistrust of teachers and the schools who are viewed as discriminatory. They view American society as not fully rewarding or accepting their education and hard work and thus are ambivalent about the notion that education is a tool that results in success or aids in overcoming barriers to upward mobility and economic success (Ogbu & Simons, 1998).

Empirical research on the other factors that contribute to the academic underperformance of minorities have focused on, family/personal and community, as well as institutional/structural variables. Family/personal and community factors include: the disproportionate representation of Black and Hispanic students from low income households (Education Information and Accountability Services, 2009; Woods, Kurtz-Costes, & Rowley, 2005); family structure that is dominated by single parent households (Bankston III & Caldas, 1998; Pong, Dronkers, & Hampden-Thompson, 2003; Puma,

Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Rathbun & West, 2004); high rates of teen pregnancy (Pong, Dronkers, & Hampden-Thompson, 2003); low family social status (Caldas & Bankston III, 1997); less stimulating home environment, particularly during the preschool and elementary years (Baharudin & Luster, 1998); highest maternal educational level at less than a high school diploma or its equivalent (Bankston III & Caldas, 1998; Liu & Lu, 2008; Rathbun & West, 2004); poor parental involvement (Lee, Bryk & Smith, 1993; Tillman, 2005); lowered parental academic expectations (Lee, Bryk & Smith, 1993; Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997); incongruence between home culture and mainstream school culture (Lovelace & Wheeler, 2006; Ogbu & Simons, 1998; Strand, 2012); lower proficiency with the English language (Haladyna, Nolen & Haas, 1991); less positive attitudes toward science (Singh, Granville, & Dika, 2002); less positive views of science ability (Sikora & Pokropek, 2012); lower levels of student motivation (Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Singh, Granville, & Dika, 2002; Strand, 2012); limited background knowledge (Fisher, Ross, & Grant, 2010); poor literacy rates (Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Reeves, 2000); lower levels of student attendance (Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Strand, 2012); student participation in less academically rigorous courses (Moore & Slate, 2008; National Science Foundation, 2012a; Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Rumberger & Palardy, 2005; Strand, 2012); lower degree of fitness and higher proportion of fatness (Chomitz, Slining, McGowan, Mitchell, Dawson, & Hacker, 2009; Davis & Cooper, 2011; Gurley-Calvez & Higginbotham, 2010; Hollar, Messiah, Lopez-Mitnik, Hollar, Almon, & Agatston, 2010);

and poor community emphasis on the importance of education (Ogbu & Simons, 1998; Price, 2005).

Institutional/structural factors include: staffing problems at schools with high minority populations (Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Tillman, 2005); shortage of qualified teachers in the classroom (Murphy, DeArmond &Guin, 2003; National Science Foundation, 2012a); lowered teacher expectations (Gregory & Weinstein, 2008; Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Rumberger & Palardy, 2005; Strand, 2012); poorer level of instruction from teachers (Desimone & Long, 2010); receipt of instruction from struggling or teachers new to the teaching field (Desimone & Long, 2010); the disproportionate representation of Black students in discipline referrals and its subsequent effect on student removal from the classroom (Aud, Fox, & KewalRamani,. 2010; Gregory & Weinstein, 2008; Skiba, Michael, Nardo, & Peterson, 2002; Strand, 2012; Vincent, Tobin, Hawken, & Frank, 2012); negative racial climate in heterogeneous schools denoted by lack of racial fairness and experiences of racism (Mattison & Aber, 2007); lower degree of safety expressed by students at school (Rumberger & Palardy, 2005); disproportionate representation of students from similar low socioeconomic standing attending schools with those from similar backgrounds (Bankston III & Caldas, 1998; Caldas & Bankston III, 1997; Condron & Roscigno, 2003; Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Rumberger & Palardy, 2005; Ryabov &Van Hook, 2007); similar ethnic composition of schools (Aud, Fox, & KewalRamani, 2010; Bankston III & Caldas, 1998; Condron & Roscigno, 2003; Lee, Bryk & Smith, 1993; Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997); disparity in educational funding in which high-

poverty districts receive less resources (Butchart, 2010; Condron & Roscigno, 2003; Chambers, Levin, & Parrish, 2006; Ryan, 1999; Thompson, Wood, & Crampton, 2008); and the negative impact of stereotypes about a group's intellectual ability or competence based on ethnicity (Steele & Aronson, 1995; Strand, 2012; Walton & Spencer, 2009; Woods, Kurtz-Costes, & Rowley, 2005).

The disparity in scores has, in part, been attributed to the concept known as stereotype threat. "Stereotype threat is conceived as a state of psychological discomfort that, if sufficiently acute, can impair performance. It is thought to arise when students are confronted with an evaluative situation, in which a stereotype regarding a particular ability is relevant" (Appel, Kronberger, & Aronson, 2011, p. 904). According to Hoy and Hoy (2009), “when stereotyped individuals are in situations where a stereotype applies, they bear an extra emotional and cognitive burden. The burden is the possibility of confirming the stereotype, in the eyes of others or in their own eyes” (p. 293).

This burden can result in anxiety and thus undermine performance (Hoy & Hoy, 2009) by negatively affecting the intellectual functioning of these students, particularly on standardized tests (Steele & Aronson, 1995) and increasing negative domain-specific thinking (Cadinu, Maass, Rosabianca, & Kiesner, 2005). If this threat continues, it may result in students disassociating with achievement in school and related intellectual domains as a protective mechanism; thus altering their self-concept in a manner in which school achievement does not serve as the basis of self-evaluation or personal identity (Steele & Aronson, 1995). The negative impact of stereotypes is not solely limited to race, however, but instead to any categorical grouping of individuals in

which a stereotype exists, including those from low socioeconomic backgrounds and of female performance in math and science.

The Current Status of Minorities in Science

Despite the paltry scores earned by Black and Hispanic students, trends in academic performance allude to the notion that there is hope as the average score of Black and Hispanic students have increased since 1995 at both the fourth and eighth grade level (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008). The caveat however, is that the average scores of both Black and Hispanic students continue to lag behind their White and Asian peers. In fact, the average scores between White and Black students in 2007 resulted in a 79 point and 96 point difference at the fourth and eighth grade level respectively, and a difference of 65 and 71 points between fourth and eighth-grade White and Hispanic students, respectively (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008). However, recent assessments have indicated that the achievement gap between the different ethnicities are narrowing (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008; National Center for Education Statistics, 2012). For example, the NAEP Science 2011 data indicates that when compared with 2009 data, the gap in scores was reduced from 36 points to 35 points between White and Black students, and from 30 points to 26 points between White and Hispanic students (National Center for Education Statistics, 2012). Therefore although an achievement gap exists between White and Asian students and their Black and Hispanic peers, student achievement for Black and Hispanic students are improving.

Based on these staggering statistics, it is not surprising that a disparity in ethnic representation in the sciences exist. Although Blacks and Hispanics comprised 12%

and 15% of the total U.S. population, they comprise smaller percentages of engineering and science degree recipients and employed scientists and engineers compared to the population; and thus are considered underrepresented in science and engineering (National Science Foundation, Division of Science Resources Statistics, 2011; National Science Foundation, 2012a). In fact, the science and engineering workforce is primarily composed of White males (National Science Foundation, Division of Science Resources Statistics, 2011; National Science Foundation, 2012a) and that demographic is also primarily represented in graduate students and post doctorates in science and engineering in 2010 (National Science Foundation, 2012b). African American and Hispanics represent substantially lower levels of participation in science and engineering and other professional and related occupations than the U.S. workforce as a whole (National Science Foundation, Division of Science Resources Statistics, 2011) although at the elementary level, White and Black males express similar levels of positive attitudes towards science (Riegle-Crumb, Moore, & Ramos-Wada, 2010). In fact, Hispanic students are more likely to attain associate degrees than bachelor's degrees (National Science Board, 2004).

However, underrepresented minority students obtaining bachelor's and master's degrees in science and engineering have increased over two decades since 1989, although those earning doctorates in these fields flattened since 2000 (National Science Foundation, Division of Science Resources Statistics, 2011). Of the science and engineering degrees earned at the bachelor's level by underrepresented minorities, the greatest rise has been in social, computer, and medical sciences fields of study (National Science Foundation, Division of Science Resources Statistics, 2011). In fact,

the percent of minorities in the doctoral science and engineering labor force increased from 4% in 1990 to 6% in 2007 (National Science Foundation, 2011).

While this increase in minority representation in the sciences represents a step forward in the right direction, significant inroads still need to be made to address student achievement disparities in light of the current demographic shift within the U.S. population. According to United States Census Bureau, Population Division (2012), minorities such as Hispanics, Blacks, and Asians represent the largest net increase in the population between 2000 and 2009. The United States Census Bureau (2012) also reported that 50.4% of the nation's population under the age of 1 was comprised of minorities, as of July 1, 2011. As of 2011, there were 114 million minorities, primarily composed of Hispanics who numbered 52 million in 2011 (United States Census Bureau, 2012). Hispanics also represented the fastest growing population with a growth rate of 3.1% since 2010; thus comprising 16.7% of the nation's total population in 2011 (United States Census Bureau, 2012). African Americans however, represented the second largest minority group at 43.9 million in 2011, an increase of 1.6% from 2010 (United States Census Bureau, 2012).

While the shift in demographics is not particularly troubling in itself, when viewed within the context of an increase in the retirement of workers in the science and engineering workforce, a reduction in White male participation in science and engineering, more foreigners attaining U.S. science and engineering degrees, particularly advanced degrees in the past 2 decades; and an increasing minority population whose participation rate is half or less than their White peers and have been traditionally underrepresented in science and engineering (National Science Board,

2004); the need for minorities pursuing degrees in science and engineering becomes of increasing importance.

Socioeconomic Status and Science Achievement

Empirical research has conclusively indicated that student achievement is negatively affected by the low socioeconomic status of students (Caldas & Bankston III, 1997; Desimone & Long, 2010; Fleischman, Hopstock, Pelczar, & Shelley, 2010; Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008; Liu & Lu, 2008; Maerten-Rivera, Myers, Lee, & Penfield, 2010; Maerten-Rivera, Myers, Lee, & Penfield, 2010; National Center for Education Statistics, 2012; National Science Foundation, 2012a; Rathbun & West, 2004). Thus the condition of poverty has greater implications that go beyond mere access to financial resources. According to the United Nations Children's Fund [UNICEF] (2012), failing to protect children from poverty is a costly mistake borne by not only children, but the nation of those children due to the reduction in skills and productivity, lower levels of health and academic achievement, increased probability of unemployment and dependence on welfare, increased judicial and social protection systems, as well as the loss of social cohesion. According to the United Nations Children's Fund [UNICEF] (2012), 23.1% of American children are living in relative poverty (defined as "living in a household in which disposable income, when adjusted for family size and composition, is less than 50% of the national median income" (p. 3). When compared internationally, the U.S. ranked as the 2nd highest on relative child poverty rates out of the 35 economically advanced countries (UNICEF, 2012).

Free and Reduced-Price Lunch

One service provided to children living in poverty is the National School Lunch Program (NSLP) which is federally assisted and operates within approximately 100,000 public and nonprofit private schools and residential child care institutions (United States Department of Agriculture, 2012). NSLP provides "nutritionally balanced, low-cost or free lunches to more than 31 million children each school day in 2011" (United States Department of Agriculture, 2012, para. 1). According to the Education Information and Accountability Services (2009), the NSLP was established in 1946 to provide free and reduced-price lunches to students from economically disadvantaged families. Children eligible for free meals are from families with incomes at or below 130 percent of the poverty level and those eligible for reduced-price meals are from families with incomes between 130 percent and 185 percent of the poverty level (United States Department of Agriculture, 2012).

In the 2006-2007 school year, 1.2 million students in Florida qualified for free and reduced-price lunch (Education Information and Accountability Services, 2009). This figure represents a 24.69% increase or an increase of 257,214 students since the 1999-2000 and 2008-2009 academic school years. Of students eligible, Black and Hispanic students represented the largest proportions of all ethnicities in which 69.59% of Black and 62.38% of Hispanic students were eligible; a significant quantity compared to the 43.98% average of all ethnicities combined (Education Information and Accountability Services, 2009).

Research on the Impact of socioeconomic on Achievement

According to Caldas and Bankston III (1997), research has consistently demonstrated since the mid-1960s that individuals with a low socioeconomic standing

are more likely to engage in lower levels of academic achievement than individuals of a higher socioeconomic standing, even while controlling for a variety of other factors. One such study by the National Center for Education Statistics (2012), categorized students as coming from lower-income families based on student eligibility for the NSLP for either free or reduced-price school lunch. The study indicated that students not eligible for free or reduced-price school lunch obtained an average science score that was 27 points higher than their lower-income peers. The study also indicated that when compared with 2009, the average scores of both eligible and non eligible students increased by 4 points for those eligible and 3 points for non-eligible students. However, no significant changes between 2009 and 2011 in the score gap between both groups occurred.

Empirical studies also have indicated that students from low socioeconomic status score lower than those from higher socioeconomic standing in combined reading literacy (Fleischman, Hopstock, Pelczar, & Shelley, 2010) and science measures (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008). Additionally, students attending the highest poverty public schools (in which 75% of students were eligible for free or reduced-price school lunch) had the lowest average science score in both fourth and eighth grade than students attending schools in which the percent of students eligible for free or reduced-price lunch was 50-74.9%, 25-49.9%, 10-24.9%, and less than 10% respectively (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008). This finding is consistent with Puma, Karweit, Price, Ricciuti, Thompson, and Vaden-Kiernan's 1997 study that indicated that the academic performance of children from high- poverty schools was lower than those in low-poverty schools. Children from low income households started school, on average,

academically behind students in low poverty schools and were unable to close the gap in achievement as they progressed through school (Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997).

Thus although several studies have highlighted the negative impact of low socioeconomic status on student achievement, other studies have alluded to a similar, yet nuanced variable. According to Easton-Brooks and Davis (2007), it is wealth, not socioeconomic status that accounts for the more significant variance in educational outcomes, particularly for African-American students. It is wealth that confers advantages beyond income, occupation, and parental education (Easton-Brooks & Davis, 2007). It provides psychological benefits as well as the financial aspects that affect the degree of personal security experienced, one's ability to finance a collegiate education, allows for a variety of options in places to live, as well as affords the protection from the stress of short-term unemployment and other emergencies (Easton-Brooks & Davis, 2007).

Caldas and Bankston III (1997) also revealed that the family social status has a more significant effect on student achievement than socioeconomic status as measured by qualification for free and reduced-price lunch. They noted that students who attended schools with high minority concentration tended to contain peers of relatively low family social status backgrounds and were marked by higher levels of poverty. They also discovered that low socioeconomic has a small, negative impact on individual student achievement. In fact, attending school with students who are from higher socioeconomic backgrounds has a positive impact on increasing student achievement, independent of one's socioeconomic background, race, and other factors.

This is because the more affluent peers tend to bring additional resources associated with higher family educational and occupational status. Additionally, students of all ethnicities are negatively affected by poverty although African American achievement is more negatively affected that their White peers. This supports what proponents for heterogeneous schools along socioeconomic lines such as Coleman, Campbell, Hobson, McPartland, Mood, Weinfield, and York (1966) have long advocated. According to them, the economic diversity of the school body serves as a means of enabling those from low socioeconomic lines to benefit from contact with those more socially advantaged as a means of acquiring "social capital."

Empirical studies have consistently noted the intersection between socioeconomic standing and ethnicity, as previously discussed in the "Ethnicity and Student Achievement" section of Chapter 2 (Aud, Fox, & KewalRamani, 2010; Caldas & Bankston III, 1997; Easton-Brooks & Davis, 2007; Snyder, Dillow, & Hoffman, 2008) as proportionately more African-Americans occupy low socioeconomic than their White counterparts (Caldas & Bankston III, 1997; Snyder, Dillow, & Hoffman, 2008). Thus, race and ethnicity are difficult to conspicuously distinguish from socioeconomic status as impacting on student achievement due to the disproportionate representation of racial or ethnic minorities in low socioeconomic settings (Maerten-Rivera, Myers, Lee, & Penfield, 2010). According to Snyder, Dillow, and Hoffman (2008), 33% of all Black families with children under 18 reside below the poverty level compared to 9.5% of Whites and 26.6% for Hispanics in 2006. The statistic is significantly higher in single-parent households lacking a husband as the poverty level for Whites with no husband increased to 32.9%, 49.7% for Blacks, and 47.2% for Hispanics (Snyder, Dillow, &

Hoffman, 2008). Thus it is not surprising that African American children are more likely to live in poverty than their peers of other ethnicities. In fact, 34% of black children live in poverty, compared to 27% Hispanics, and 10% Whites (Aud, Fox, & KewalRamani, 2010). Additionally, of the 48% of fourth graders in public school that were eligible for free-reduced price lunch in 2009, 77% were Hispanic, 74% Black, and 29% White (Aud, Fox, & KewalRamani, 2010). Therefore, since ethnicity and socioeconomic status are intertwined, some of the some of the cited reasons for student underperformance in Chapter 2 under the subsection, "Ethnicity and Science Achievement" applies.

Research has also indicated the intersection between socioeconomic standing and the number of parents in the household (Puma, Karweit, Price, Ricciuti, Thompson, Vaden-Kiernan, 1997; Rathbun & West, 2004). Research has indicated that 55% of children living in poverty reside in single-parent households (Rathbun & West, 2004). One particular study, Rathbum and West's 2004 study used data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) that examined specific risk factors on student achievement. These risk factors included living in a single-parent household, living below the federal poverty level, living in a household whose primary home language is other than English and the highest level of maternal education less than or below a high school diploma or its equivalent.

The results indicated that as the number of risk factors of a child's household increase, student gains between the start of kindergarten and the end of third grade decrease in reading and math when compared with children with fewer family risk factors, after controlling for other child, family, and school characteristics. Furthermore, students in such households are less likely to be proficient in reading, mathematics, and

science than those with less risk factors. Children with no or low associated risk factors scored significantly higher than those with more risk factors. Puma, Karweit, Price, Ricciuti, Thompson, and Vaden-Kiernan's 1997 study also demonstrated that students performing poorly academically were "participants [that]were more economically disadvantaged. They were more likely to be living with a single parent, have a total family income under $10,000, receive public assistance, and have poorly educated parents whose natural language is not English. They were more likely to be non-White and live in urban areas with the concomitant increased likelihood of being exposed to high rates of crime, physical violence, drug abuse, and substandard living conditions" (p.22).

Thus the impact of the environment is a significant one in the development of the cognitive capabilities of the child. One such environment that has received attention is the verbal environment during the early stages of child development. According to Christensen, Horn, and Johnson (2011), “a significant portion of a person’s intellectual capacity is determined in his or her first 36 months” (p. 149) due to exposure to verbal interactions. Merely speaking with a child was however not indicated to improve intellectual capacity but instead, research indicates that children in pre-K need significant exposure to “language dancing.” “Language dancing” involves face to face discussions with the infant in which the parent speaks in a fully adult, sophisticated, chatty language. It is characterized by deliberate, uncompromised, personal adult conversation. Limiting conversations to “business” in which statements are made about what needs to be done are often characterized by simple, direct, here-and-now conversations.

Thus their ability to improve cognitive development is limited. According to Zegiob and Forehand (1975), the socioeconomic status of the mother was a significant determinant of maternal behavior. The study found that less directive and more social interchange, especially verbally, were common among middle-class mothers than their lower-class counterparts. Lower class mothers also gave more commands than middle-class mothers and they were more directive and controlling. Also, less verbal exchanges occurred between lower-class mothers and their children. Tulkin and Kagan's 1972 study also examined how the different social class backgrounds of individuals affected the experiences of infants. The study also concluded that middle-class mothers vocalized every verbal behavior more frequently than lower-class mothers. Thus by improving the linguistic environment of the home, lower-class mothers can help to improve the cognitive development of their children as a means of improving student academic performance in the future.

Research has also indicated that poverty and being a minority does not have to negatively impact on student performance (Reeves, 2000). Reeves's 2000 study revealed a low (-0.2) correlation between student performance and being poor. In this study 90/90/90 schools were examined- that is, schools in which approximately 90% of the school's composition is ethnic minorities, a minimum of 90% are eligible for free and reduced-price lunch, and approximately 90% of its students are achieving at proficiency levels based on independently conducted tests of student academic achievement. Reeves (2000) realized that these schools focused on a limited number (in particular 5) areas in which to improve and made curriculum choices whereby students spent more time in core subjects such as reading, writing, and mathematics and less time on other

subjects. Consequently, although significant instruction was lacking in science, these students nevertheless outperformed their peers of similar backgrounds thus highlighting that reading and writing are critical in determining the academic achievement of students. Furthermore, they changed the school schedule so that at the elementary level students were engaged in 3 hours of literacy and at the secondary level they provided double periods of English and mathematics (Reeves, 2000).

Gender and Science Achievement

Research has consistently demonstrated that female performance in science continues to lag behind their male peers as measured through various scientific assessments (Florida Department of Education, 2012b; Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008; Maerten-Rivera, Myers, Lee, & Penfield, 2010; O'Reilly & McNamara, 2007; National Center for Education Statistics, 2012; Riegle-Crumb, Moore, & Ramos-Wada, 2010; Shymansky, Yore, & Anderson, 2004). According to National Center for Education Statistics (2012), male students scored higher than females in science. The average male science score was higher than those of females in grade four by 5 points (although not statistically significant) and increased in grade 8 to 12 points (statistically significant at $p < 0.05$). Although there was no significant difference in science performance of male and female fourth-graders, males outperformed females by 5 points only in Earth Sciences, thus excluding all other science content domains such as Life Sciences and Physical Sciences (National Center for Education Statistics, 2012).

In the eighth-grade however, males outperformed females in Biology, Physics, and Earth Sciences ($p<0.05$). When the 2003 and 2007 average scores were analyzed, data indicated that the difference in scores between males and females decreased from

12 scale score points to 5 scale score points for fourth-graders. At the eighth grade level however, no significant difference in changes in the gap between average science scores from 1995 and 2007, thus maintaining the male advantage in average science scores (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008). Considering that individuals in science-related fields earn considerably more on average than other fields (National Science Foundation, 2011), women are at a disadvantage in future earnings due to their decreased proficiency levels in the sciences.

However, a study of Iranian secondary level girls and boys revealed otherwise (Nasr & Soltani, 2011). According to Nasr and Soltani (2011), Iranian secondary level girls tend to have better achievements than their male counterparts. While some may account for this difference due to discrepancies in student attitudes based on gender, the study indicated no significant difference on attitudes towards biology between secondary level girls and boys (Nasr & Soltani, 2011). Additionally, according to Nasr and Soltani (2011), there is no significant difference between attitude towards biology and students' achievement in biology courses ($p<0.05$, $r = 0.12$). However, of the many attitudes towards biology dimensions studied, the only positive and significant dimension was "biology is fun for me."

George (2006) found contradictory results, revealing an initial gender difference in the attitude toward science and its utility. According to George (2006), in the 7th grade, boys have a more favorable attitude toward science and its utility than girls. Bhanot and Jovanovic (2009) also revealed that boys were more confident in their science abilities although no gender differences existed in actual science ability or science performance. Greenfield (1997) on the other hand, failed to find any gender

differences in attitudes. According to Greenfield (1997), elementary girl students liked science more than elementary boys, and although attitudes for both genders declined during middle school years, high school boys' attitudes rose again but not the girls. Thus, when all grades are combined, no significant gender difference in attitude exists (Greenfield, 1997). George (2006) also noted that the attitude of boys changes faster than girls as they advance in coursework. Furthermore, a higher proportion of students expressed a more favorable attitude toward the utility of science during their 10th grade while taking biology and lower favorability while taking earth or physical science (George, 2006).

Prokop, Prokop, and Tunnicliffe (2007) also revealed no significant gender differences in preferences toward biology lessons although Slovakian girls reported higher interest in biology than Slovakian boys ($p<0.05$). Boys also considered biology lessons to be more difficult than girls ($p<0.05$) while the mean score for girls reporting that biology lessons were more important was significantly higher than in boys ($p<0.001$). Biology was also the preferred science course for 4th and 6th graders (ages 9-10 and 11-12) though, younger (between ages 6-8) and older (between ages 13-15) students reported lower rates of course preference (Prokop, Prokop, & Tunnicliffe, 2007). Thus, despite the trend of the changing reports of the utility of science based on the science subject matter being learned, attitude towards science declines during the middle and high school years (George, 2006; Greenfield, 1997) although attitudes regarding its utility is positive (George, 2006). This waning positive attitude toward science as students advance in science coursework is partly responsible for the decline in science learning (as evident in scores) and interest in science-related careers

(George, 2006). Riegle-Crumb, Moore, and Ramos-Wada (2010) also revealed that enjoyment of science serves as an important factor in pursuing science careers specifically for White and Hispanic females although achievement in science served as a more significant influence for Black students at younger ages.

Proposed Explanations for Disparity in Achievement

Research has documented that attitudes and achievement among adolescents in science is influenced by a myriad of factors (Taltont & Simpson, 1985). Factors include parental background and family environment, individual characteristics such as locus of control, achievement motivation, and self-concept; as well as the influence of the school such as teacher and administrative styles as well as class climate (Taltont & Simpson, 1985). For example, Hidi and Harackiewicz's 2000 study that revealed that learning and academic performance is strongly influenced by interest, goals, and motivation. Velayutham, Aldridge, and Fraser (2012) revealed that motivational beliefs regarding learning goal orientation, task value, and self-efficacy influenced student learning. Research also indicated that student self-concept is also a strong predictor of attitude towards science and its utility (George, 2006) and males posses a higher self-concept in science than females (Sikora & Pokropek, 2012). Additionally, teachers who encouraged students to pursue opportunities in science and work hard in it were more likely to have students with positive attitudes towards the field (George, 2006). Friends also play a role in influencing course enrollment. Riegle-Crumb, Farkas, and Muller (2006) revealed that females with female friends who excelled academically in both traditionally male and female subjects positively affected advanced course taking in those subjects.

Research has also exposed the effect of stereotype threat on female performance in math and the sciences. Considering that the gender gap in performance does not emerge until high school (National Center for Education Statistics, 2012), gender differences in ability and preference may not represent sex differences due to genetic, chemical, or biological factors (Good, Woodzicka, & Wingfield, 2010), but instead learned through socialization (Bhanot & Jovanovic, 2009; Bleeker & Jacobs, 2004; Cadinu, Maass, Rosabianca, & Kiesner, 2005; Walton & Spencer, 2009). Research has indicated that regardless of age or gender, implicit stereotyping of women in science predicted the sex differences in performance on the TIMMS assessment of eighth graders in 1999 and 2003 (Nosek, Smyth, Sriram, Lindner, Devos, Ayala, Bar-Anan, Bergh, Cai, Gonsalkorale, Kesebir, Maliszewski, Neto, Olli, Park, Schnabel, Shiomura, Tudor Tulbure, Wiers, Somogyi, Akrami, Ekehammar, Vianello, Banaji, & Greenwald, 2009).

Haworth, Dale, and Plomin (2009) reported that genetic and environmental influences are important for high science performance in both genders but the effect of the environment becomes increasingly more influential with age. Additionally, Bhanot and Jovanovic (2009) revealed that the science ability of boys were overestimated by their parents than parents of girls, thus partly explaining the tendency of girls to undervalue their science ability and boys in overestimating it. Furthermore, parents of boys perceived that their sons liked science more than parents of girls (Bhanot & Jovanovic, 2009). However, when the mother or father of a girl believed in the importance of science, girls valued science more although parents played no role in influencing the beliefs of boys (Bhanot & Jovanovic, 2009).

Informal discussions by mothers regarding the importance of science and possible careers in science also served in the creation of positive attitudes by girls (Bhanot & Jovanovic, 2009; Dewitt, Archer, Osborne, Dillon, Willis, & Wong, 2011). Notable however, is the finding that mothers are more likely to encourage science interest in their sons who are not doing well and only become involved with girls when they fare well academically (Bhanot & Jovanovic, 2009). Bleeker and Jacobs (2004) and Dewitt, Archer, Osborne, Dillon, Willis, & Wong (2011) also revealed that the self-efficacy of adolescents in math and science as well as their career choices were related to their mother's beliefs about their adolescents' abilities. Thus early maternal expectations have an enduring effect on children and their career decisions later in life as children are "taught" which academic pursuits are considered appropriate based on their gender and learn role expectations which may lead to the gender gap (Bleeker & Jacobs, 2004).

Research has also indicated that the school context may also serve as a contributor to the effects of stereotype threat, even in Non-Western locations (Pico & Stephens, 2012). Women attending single-sex schools exhibit higher levels of motivation, domain identification, self-efficacy, and performance in male dominated domains such as mathematics than those attending coed institutions (Pico & Stephens, 2012). Even when stereotypes were applied, attending a single-sex school served as a buffer, thus failing to negatively affect performance, although those attending coed institutions reported significantly worse performance (Pico & Stephens, 2012). The proposed difference in the impact on student academic performance as a factor of gender composition of the school is due to the salience of stereotypes in the

environment. According to Pico and Stephens (2012), environments in which stereotypes are salient serve as grounds for stereotypes to be triggered, thus exerting its impact. Thus the chronic exposure of females to stereotype laden environments results in increased susceptibility to its effects by (1) increasing accessibility and potential believability of gender stereotypes, and (2) serving as an incubator in which girls adapt to dispositional factors such as stigma consciousness in such contexts over a period of time; thus catalyzing stereotype threat and subsequently degrading performance (Pico & Stephens, 2012).

Unfortunately, science textbooks have also served to perpetuate gender stereotypes via the images and languages utilized (Good, Woodzicka, & Wingfield, 2010). According to Good, Woodzicka, and Wingfield (2010), "girls may begin to believe in the stereotype, attributing their inferior performance to a supposed natural sex difference, and ultimately leading them to avoid majors and professions involving math and science. Thus, textbooks, considered positive instruments of learning, may ironically teach girls that they have no place in the academic areas of math and science and thereby reduce their achievement and enjoyment within these disciplines" (p. 145). Although the authors do not advocate that the elimination of biased textbook images will assuage stereotype threat, they instead advocate that the gender gap in science performance will be lessened (Good, Woodzicka, & Wingfield, 2010). In fact, a study conducted by Appel, Kronberger, and Aronson (2011) revealed that stereotype threat even impairs the ability to learn and acquire knowledge in preparation for a test which serves as a precondition for the potential occurrence of stereotype threat during testing.

According to their study, women who were aware of the stereotype that women were less proficient in science, technology, engineering, and mathematics (STEM) test preparation than men were impaired in their note-taking activities as well as evaluating the notes of others, especially if they identified with the female gender (Appel, Kronberger, & Aronson, 2011). According to Appel, Kronberger, and Aronson (2011), "if stereotype threat also impairs learning activities (at least among those who are domain identified), then, over time, targets not only will demonstrate impaired test performance but will actually learn content in less efficient ways as well. Gradually, the knowledge gaps between targets and nontargets will widen" (p. 911).

During the learning process as well, stereotype threat surfaces in other ways. There are reports of female students not receiving as much teacher attention in science classes although they initiated as many teacher interactions as their male peers (Greenfield, 1997). During the test-taking process, research has shown that females perform better academically in science when exposed to counter-stereotypic images than stereotypic images (Good, Woodzicka, & Wingfield, 2010). Thus, the behaviors that develop student capabilities are negatively affected by stereotypes, thus underscoring the importance of gender-neutral textbooks, language, and activities within the classroom. Thankfully, the developers of the Biology I end-of-course assessment have intentionally ensured that test questions are equal in its specification of gender and ethnicity (Florida Department of Education, 2011c) as one means of reducing the impact of stereotype threat on student performance.

According to American Association for the Advancement of Science (AAAS) Project 2061 (1990), when the science education of any student (particularly girls and

minorities) are neglected, students are deprived of a basic education which handicaps them for life as well as deprives the nation of a talented workforce and informed citizens; a loss that the nation cannot afford.

Representation of Women in Science

Historically, women have been underrepresented in the scientific fields and scientific careers (National Science Foundation, 2012a). According to Long and Fox (1995), the institution of science is marred with inequality in career attainments whereby women and most minorities, as groups, participate at lower levels, have lower positions, lower productivity levels, and less recognition than White men. Thankfully, this male-stereotypical view of science and scientists is less a part of the female consciousness than males (Greenfield, 1997). Despite the stereotypical male view of science and scientists, there was a significant difference in GPA between men and women in any field of the sciences, in which women had higher GPAs than their male counterparts (Sonnert & Fox, 2012) although they were more likely to underestimate their science ability (Sikora & Pokropek, 2012). Additionally, according to Sonnert, Fox, and Adkins, (2007), the percentage of women majors enrolled in majors in the scientific and engineering fields has increased steadily in a linear manner over a 16- year period (between 1984 and 2000).

In fact, since 2004, females have attained more doctoral degrees in science and engineering than males (National Science Foundation, 2011). Moreover, in 2010, more women than men were conferred with scientific degrees (National Science Foundation, 2012b) and the percent of women holding doctoral degrees in science and engineering has increased from 23% of the science and engineering doctoral labor force in 1990 to 34% in 2007 (National Science Foundation, 2011). Even more so, more women than

men enroll in college and graduate with a bachelor's degree irrespective of all racial/ethnic groups (National Science Foundation, Division of Science Resources Statistics, 2011).

Despite this trend of increasing women graduating with undergraduate degrees, a higher proportion of degrees in science and engineering fields of study were earned by proportionately more males (National Science Foundation, 2011; National Science Foundation, Division of Science Resources Statistics, 2011; National Science Foundation, 2012a) which is not alarming considering that at the elementary level, White, Hispanic, and Black females reported significantly lower rates of interest in pursuing science careers than their White male counterparts and also reported lower levels of science enjoyment and science self-concept overall (Riegle-Crumb, Moore, & Ramos-Wada, 2010). At the secondary level, although more women are enrolled in Advanced Placement courses than their male counterparts, fewer women are enrolled in science courses and experienced less success, as a lower percentage scored at or above the criterion level in those subjects (Moore & Slate, 2008). Consequently, there is an uneven distribution of women in the scientific fields. According to National Science Foundation (2011) and Sonnert and Fox (2012) women have a higher percentage of representation in life sciences. The participation of women among other scientific fields also varies although within fields it has remained relatively consistent over every level of degree.

For example, women participate at higher levels in psychology and medical sciences. Women's participation in biosciences has also increased at all degree levels although at the master's and bachelor's level, the growing trend has stabilized over the

last 5 years (National Science Foundation, Division of Science Resources Statistics, 2011). At the secondary level, Bleeker and Jacobs (2004) and Sikora and Pokropek (2012) also reported that the life sciences were selected by more females interested in pursuing careers in the sciences than other scientific fields such as physical sciences. When compared with their male counterparts in mathematics and physical sciences, women earn degrees at medium to low levels (National Science Foundation, Division of Science Resources Statistics, 2011) as more than 80% of bachelor's degrees in computer science, physics, and engineering were earned by males in 2007 (National Science Foundation, 2011; National Science Foundation, 2012a).

Therefore, research has indicated the increasing participation and achievement of women in the Biological sciences, which stimulates discussion on whether a gender gap even exists within biology. National Center for Education Statistics (2012), indicated that males outperformed females in biology at the eighth-grade level, but the gap in scores decreased dramatically among fourth-graders to only 5 points. Those fourth-graders in 2007 would represent the cohort of ninth-graders in this study, if the traditional educational trajectory is followed. Thus, the possibility exists that these students may represent a cohort in which gender is negligible in student achievement in biology. This does not imply that gender differences may no longer exist among other sciences, but instead may not be applicable to biology. Therefore, the context of this study will test such an assertion.

Summary

The purpose of this study is to determine the influence and relationship of four variables- FCAT Reading level, ethnicity, socioeconomic status, and gender- on biology achievement as reflected on the results from the Florida Biology I EOC assessment at a

Florida Title 1 high school. Results from previous studies have indicated that achievement in science is differentially affected by each of the four variables. Research has consistently indicated that students with low levels of reading proficiency score significantly lower on science assessments than their peers who are proficient in reading. Research also indicates that students from minority backgrounds and students who are eligible for free and reduced-price lunch score lower than students who are White and who are not eligible for free and reduced-price lunch. Last, research has also indicated that women score lower on science assessments than their male counterparts. The proposed explanations that account for these differences reveal that this phenomenon of underperformance is a complex one and interdependent on several other factors. Consequently there are significant implications not only on graduation rates but also on the ability of the education sector to prepare students for a more technologically advanced and global society.

CHAPTER 3
METHODOLOGY

This chapter will present an overview of the quantitative methods utilized in this study. The subsections that comprise this chapter include research design, population, sample and sampling procedures, procedure for data collection, measures, participants, and method of data analysis.

Research Design

This study adopted a causal-comparative exploratory research or *ex-post facto* design. According to Gay, Mills, and Airasian (2009), causal-comparative research describes preexisting conditions and therefore is considered descriptive research. However, this methodology also seeks to determine the causes or reasons for the existing differences in the status of groups. Thus, the researcher aims to identify the major factor that has led to the difference between groups as "both the effect and the alleged cause have already occurred and must be studied in retrospect (Gay, Mills, & Airasian, 2009, p. 218). This methodology was deemed appropriate as the study collected preexisting data from the sample being investigated, and therefore given the *ex-post* facto design; the researcher did not consciously or deliberately manipulate any of the variables of interest in the study. The researcher also did not deliberately control variables, including confounding variables, and failed to utilize random sampling. However, an analysis of covariance will be utilized in the study in an attempt to control for extraneous variables.

Population

Originally, the researcher wanted to utilize data obtained from two schools for the study. However, due to the extensive data collection protocols that required

considerable time, significant delays in the study would have occurred. Consequently, the researcher elected to focus on Title 1 school as the focus of this study. Thus, the target population for this study included the high school students matriculated at a rural North Florida Title 1 school. The school was purposely sampled as it is a co-educational institution whose demographics contain relatively equal proportions of males and females (52% males and 48% females) enrolled in Biology I (Florida Department of Education, 2012, o, p), and is reflective of the major ethnic groups (White, Black, and Hispanic) represented within the state of Florida. The ethnic composition of the student body is comprised of 41.7% White, 54.3% Black, 2.3% Hispanic, 0.3% Asian, 0.9% American Indian or Alaskan native, and 0.4% two or more races (Florida Department of Education, 2012p). Thus, the minority rate is 59%.

Furthermore, the high school is comprised of 683 students (Florida Department of Education, 2012p), of which 64% of the students qualified for free and reduced-price lunch, thus the North Florida school is categorized as a high poverty or a Title 1 school. Academically, 39% of students scored at proficiency or higher (3 or higher) in reading on the state mandated assessment and only 26% of students obtained proficiency on the Science FCAT in 2011 (Florida Department of Education, 2011a), the highest reported percentage since the inception of the Science FCAT. Last, the overall school grade has varied throughout the years and consequently the accountability penalties or financial rewards. The school grade since 2000 has varied from attaining an A to earning Cs, Ds, and Fs (Florida Department of Education, 2012n).

The administrative team is composed of three males that each have attained a Master's degree and have between 7 to 14 years of experience as an administrator

(Florida Department of Education, 2011d). Of the 41 instructional staff present, 19.5% had 1-5 years of experience teaching, 39% had 6-14 years, and 41.5% had 15 or more years of experience (Florida Department of Education, 2011d). Of the instructional staff, 28.6% held advanced degrees (Florida Department of Education, 2012p), 99.3% were considered highly qualified under NCLB (Florida Department of Education, 2012p), 12.2% were endorsed in reading, 15% were newly hired (Florida Department of Education, 2012p), and 4.9% were National Board Certified teachers (Florida Department of Education, 2011d).

Within the district, 54% of students were male, 37.5% were White, 56% African American, 4.4% Hispanic, 0.2% Asian, and 78.2% were economically disadvantaged (Florida Department of Education, 2012p). Of the 2,715 students within the district, 2,049 students were eligible for free and reduced-price lunch in 2008-2009 or 75.47% of the district student population (Education Information and Accountability Services, 2009). When compared with the 1999- 2000 data, a 15.11% increase in eligibility occurred during the 2008-2009 school year (Education Information and Accountability Services, 2009). Additionally, newly hired instructional staff comprised 26.9% of the district, 29% of the instructional staff within the district contained advanced degrees, and 80.9% of classes are taught by teachers in-field (Florida Department of Education, 2012p).

Sample and Sampling Procedures

This research adopted convenience sampling due to the relatively easy accessibility for data collection but was purposely selected because the school under investigation possessed unique characteristics in terms of the geographic location, demographic composition of the student population, and the academic performance of

students historically. Of the students who completed the Biology I EOC assessment, all students were sampled.

Procedure for Data Collection

After having received IRB approval (see Appendix C) and permission from the appropriate school district personnel (see Appendix D), student data was collected at the school site level. Specific information from the 2011-2012 administration of the Florida Biology I EOC exams was limited to Biology I EOC T-scale scores and Biology I student EOC performance comparable to the state. FCAT Reading Retakes Developmental scale scores, FCAT Reading Retakes Achievement levels, FCAT Reading 2.0 Developmental scale scores, FCAT Reading 2.0 Achievement levels, student ethnicity, eligibility for free and reduced-price lunch, gender, and student grade level were also collected. No individual student names or personal information was collected. Researcher records of student data were then compiled on a computer and a paper copy was also kept in a secure location. Once all student data was compiled, data was then coded and entered into the software program, SPSS. Twenty-one (21) students with missing information from the North Florida student population were then removed from the sample.

Measures

Reading Proficiency

Reading proficiency was assessed using developmental scale scores from the FCAT Reading Retakes for 11th grade students and grade-level specific FCAT Reading 2.0 for 9th and 10th grade students.

Socioeconomic Status

Socioeconomic status was measured by eligibility for free and reduced-price lunch.

Biology Proficiency

Biology proficiency was assessed using T-scores obtained from the Spring 2012 administration of the Biology I end-of-course assessment. Students that were categorized in the top third by the state were considered proficient or obtaining a passing score in biology in this study.

Participants

The study was conducted in a North Florida school district that contains one Title 1 high school. Participants of this study were comprised of 43 ninth grade, 130 tenth grade, and 5 eleventh grade students; for a total of 178 students. Additionally, of the 178 students, 48% qualified for free and reduced-price lunch (38.8% free and 9.6% reduced), 59% were White, 37% were Black, 2.2% were Hispanic, and 1.7% were American Indian or Alaskan native. 54 % were female, and 55% scored 3 or higher on the FCAT Reading (30.3% level 3, 20.2% level 4, and 4.55% level 5). Table 3-1 represents the results of student performance by grade level from the Title 1 school population. Table 3-2 represents the sample of academic performance of students by grade level in the study.

Method of Data Analysis

A combination of descriptive statistics and other inferential statistical procedures such as a correlation (Pearson *r*), one-way ANOVAs, ANCOVAs, multiple- regression analysis, and independent samples one-tailed *t* tests, were utilized in the analysis of student data. Each appropriate test was set at a significance level of 0.05.

The researcher used a multiple-regression process to explore the relationship of predictive variables (reading proficiency, ethnicity/race, socioeconomic status, gender, and grade level) as they relate to the dependent variable (biology T- scale scores) in the study. Due to the categorical nature of some of the independent variables, coding was used to create vectors to represent the levels of categorical independent variable. The number of vectors needed for each independent variable is one less than the number of levels in the categorical independent variable. Therefore, for example, if there were 4 levels of an independent variable, 3 vectors were created. Then dummy coding was used in which the values entered in the vectors are 1s and 0s. Each level was represented by one vector, either 0 or 1. The 4 categorical variables that were dummy coded included gender, eligibility for free and reduced-price lunch, ethnicity, and grade level. Appendix D represents the coding and dummy coding of variables in the study.

In order to initially determine the predictive variables that had a statistically significant relationship to biology T- scale scores, a stepwise regression procedure was utilized. This procedure was used to fine-tune the model by providing the data needed to determine the removal of variables that did not significantly contribute to biology achievement. Based on the initial data, variables were then added or removed to determine the best model that had the largest *R* squared and adjusted *R*-squared values, which was statistically significant, and removed highly correlated predictors. The statistical significance and relative importance of each predictive variable was checked by examining the unstandardized coefficient beta weights and the standardized beta weights of each predictive variable.

Scatterplots and histograms based on the data were created as well as correlation matrices, multicollinearity statistics, and a simultaneous regression analysis using all variables. A histogram of the residuals was examined to determine whether the data is relatively normally distributed. A normal P-P plot was also used to determine how well a specific distribution fits the observed data. An analysis of the scatterplots were then conducted to determine the strength of the linear line or whether the scatterplots were unrelated to the dependent variable. Furthermore, a scatterplot of the regression standardized predicted value and regression standardized residual was used to determine whether homoscedasticity is present in the data.

Table 3-1. Student performance by grade level from the Title 1 school population

Grade level	Number of Students	Mean Scale Score	Percent scoring in the bottom third	Percent scoring in the middle third	Percent scoring in the top third
9	43	57	12	19	70
10	140	46	47	30	23
11	16	40	88	6	6
All grades	199	48	42	26	31

Table 3-2. Sample of student performance by grade level in the study

Grade level	Number of Students	Mean Scale Score	Percent scoring in the bottom third	Percent scoring in the middle third	Percent scoring in the top third
9	43	56.8	11.6	18.6	69.8
10	130	47.1	44.6	30.8	24.6
11	5	38.0	80.0	20.0	0.0
All grades	178	49.2	37.6	27.5	34.8

CHAPTER 4
RESULTS AND ANALYSIS

As previously discussed in Chapter 2, a myriad of factors contribute to the performance of students in science. The purpose of this study was to determine the influence and relationship of four variables- reading proficiency, ethnicity, socioeconomic status, and gender; on biology achievement as reflected on the results from the Florida Biology I EOC assessment at a Title 1 school.

The following paragraphs will include a restatement of the research questions and hypotheses, and the corresponding statistical analysis with further interpretation. At the conclusion of the chapter, a summary of the answers to each question will be provided. Figures 4-1 to 4-4 and tables 4-1 to 4-25 are also provided at the end of the chapter with a more detailed or visual depiction of the results.

Research Questions

- **Research Question 1.** Is there a significant difference in student performance on the end-of-course assessment in biology associated with reading proficiency? After adjusting for student gender, race/ethnicity, and socioeconomic status, is there a significant difference in student performance on the end-of -course assessment in biology associated with reading proficiency?

It was hypothesized that there will be a significant difference in student performance on the end-of -course assessment in biology associated with reading proficiency. It is hypothesized that students that are more proficient readers will score significantly higher than their less proficient peers on the Biology I EOC assessment. Furthermore, more proficient students will score significantly higher that their less proficient peers on the Biology I EOC assessment after adjusting for student gender, race/ethnicity, and socioeconomic status.

Descriptive statistics of student performance on the FCAT Reading produced the following results: the mean FCAT Reading developmental scale score was 241.01 (*SD* = 20.11), with a minimum score of 188 and a maximum score of 302. 55% of students scored a 3 or higher on the FCAT Reading assessment. Table 4-1 summarizes the FCAT reading levels of students by grade.

With alpha equal to 0.05, there was a significant relationship found between FCAT Reading developmental scale scores and Biology I EOC T-scale scores, r (178) = 0.73, $p < 0.05$. Figure 4-1 illustrates the strong positive linear relationship between FCAT Reading developmental scale scores and Biology I EOC T-scale scores. Mean Biology I EOC T-scale scores was 56.49 (*SD* = 8.78) for students who scored 3 or higher on FCAT Reading and 40.25 (*SD* = 8.99) for students who scored as a level 1 or 2 on FCAT Reading. With alpha equal to 0.05, a one tailed independent samples *t*-test indicated a significantly higher Biology I EOC T-scale score for students scoring 3 or higher on FCAT Reading compared to students scoring less than 3 on FCAT Reading, $t(176) = 12.15$, $p < 0.05$.

Table 4-2 provides descriptive statistics on the influence of FCAT Reading level on Biology I EOC T-scale scores. Mean Biology I EOC T-scale scores for students scoring level 1, level 2, level 3, level 4, and level 5 were 37.39 (*SD* = 7.88), 42.83 (*SD* = 9.23), 52.96 (*SD* = 8.24), 59.50 (*SD* = 7.37), and 66.75 (*SD* = 4.37), respectively. With alpha equal to 0.05, a one factor between subjects ANOVA indicated a significant effect of reading achievement and Biology EOC T-scale scores, F (4, 173) = 53.16, $p < 0.05$. Table 4-3 provides additional data on the results of the one-way ANOVA on FCAT reading levels and Biology I EOC T-scale scores.

Tukey post-hoc comparisons of the five FCAT Reading levels (1, 2, 3, 4, and 5) indicate that the students who attained a level 5 on FCAT Reading (M = 66.75, 95% CI [63.10, 70.40]) scored significantly higher on the Biology I EOC than students who attained a level 3 (M = 52.96, 95% CI [50.71, 55.21]), $p < 0.05$, level 2 (M = 42.83, 95% CI [39.96, 45.71]), $p < 0.05$, and level 1 on FCAT Reading (M = 37.39, 95% CI [34.81, 39.98]), $p < 0.05$. Comparisons between students who attained a level 5 and level 4 on the FCAT Reading failed to indicate a statistically significant difference, $p > 0.05$. Students who attained a level 4 on FCAT Reading also scored significantly higher on the Biology I EOC than students who attained a level 3, level 2, and level 1, $p < 0.05$. Students who attained a level 3 on FCAT Reading also scored significantly higher on the Biology I EOC than students who attained a level 2 and level 1, $p < 0.05$. Additionally, students who attained a level 2 on FCAT Reading also scored significantly higher on the Biology I EOC than students who attained a level 1, $p < 0.05$. Table 4-4 provides additional data on the Tukey post hoc analysis of the influence of FCAT Reading developmental scale scores on Biology I EOC T-scale scores.

The analysis of covariance (ANCOVA) was used to investigate the hypothesis that the observed difference in mean scores of the Biology I EOC T-scale scores is influenced by the differences in FCAT reading levels, with the covariates, ethnicity, socioeconomic status, and gender. The ANCOVA found that the means for students of varying FCAT reading levels was significant, (F = 29.87, $p < 0.05$) - that is, after the effect of the ethnicity, socioeconomic status, and gender has been accounted for. Table 4-5 provides the results of an ANCOVA for FCAT Reading level and Biology I EOC T-scale scores. Based upon the findings; hypothesis one was supported.

A further analysis of FCAT Reading developmental score data revealed interesting insights. Mean FCAT Reading developmental scale scores for 105 White, 66 Black, 4 Hispanic, and 3 American Indian students were 249.17 (SD = 16.93), 228.05 (SD = 18.65), 237.25 (SD = 15.13), and 245.33 (SD = 11.59), respectively. With alpha equal to 0.05, a one factor between subjects ANOVA indicated a significant effect of ethnicity and FCAT reading developmental scale scores, F (3, 174) = 19.78, $p < 0.05$. Tukey post-hoc comparisons of the four groups (White, Black, Hispanic, and American Indian or Alaskan native) indicate that the White students (M = 249.17, 95% Cl [245.90, 252.45]) scored significantly higher on the FCAT Reading assessment than Black students (M = 228.05, 95% Cl [223.46, 232.63]), $p < 0.05$. Comparisons between the Hispanic (M = 237.25, 95% Cl [213.17, 261.33]) and the other three groups were not statistically significant, $p > 0.05$. Comparisons between the American Indian or Alaskan native (M = 245.33, 95% Cl [216.54, 274.13]) and the other three groups were not statistically significant, $p > 0.05$. Table 4-6 presents the results of the Tukey post hoc analysis of the influence of ethnicity on FCAT Reading developmental scale scores.

The analysis of covariance (ANCOVA) was used to investigate the hypothesis that the observed difference in mean scores on the FCAT Reading is influenced by the differences in ethnicity, with the covariates, socioeconomic status and gender. The ANCOVA, found that the means for students of varying ethnicities was significant, (F = 9.0, $p < 0.05$) - that is, after the effect of the socioeconomic status and gender has been accounted for. Table 4-7 provides additional ANCOVA results for ethnicity and FCAT Reading developmental scale scores.

With alpha equal to 0.05, a one factor between subjects ANOVA indicated a significant effect of free/reduced-price lunch eligibility and FCAT reading developmental scale scores, F (2, 175) = 21.16, MSE = 329.33, $p < 0.05$. Tukey post-hoc comparisons of the three groups (free, reduced, and not eligible) indicate that the students not eligible for free or reduced-price lunch (M = 249.16, 95% CI [245.51, 252.82]) scored significantly higher on FCAT Reading than students who were eligible for free lunch (M = 230.38, 95% CI [225.77, 234.99]), $p < 0.05$. Comparisons between the reduced-price lunch (M = 240, 95% CI [231.66, 248.34]) and the other two groups were not statistically significant, $p > 0.05$. Table 4-8 provides the data for the Tukey post hoc analysis of the influence of socioeconomic status on FCAT Reading developmental scale scores.

The analysis of covariance (ANCOVA) was used to investigate the hypothesis that the observed difference in mean scores of the FCAT reading developmental scale scores is influenced by the differences in socioeconomic status, with the covariates, ethnicity and gender. The ANCOVA found that the means for students eligible for free or reduced-price lunch and those ineligible are significant, (F = 11.67, $p < 0.05$) - that is, after the effect of ethnicity and gender has been accounted for. Table 4-9 represents the results of the ANCOVA for socioeconomic status and FCAT Reading developmental scale scores.

Mean FCAT reading developmental scale scores was 242.46 (SD = 18.56) for 96 females and 239.30 (SD = 21.77) for 82 males. With alpha equal to 0.05, a one tailed independent samples t-test indicated no significant effect of gender on FCAT reading developmental scale scores, t (176) = -1.043, $p > 0.05$. Table 4-10 represents the results of the independent samples t-test to determine the influence of gender on FCAT

Reading developmental scale scores. Based upon these findings, ethnicity and socioeconomic status of a student have a significant influence on FCAT reading developmental scale scores.

- **Research Question 2**. Is there a significant difference in student performance on the end-of-course assessment in biology associated with race/ethnicity? After adjusting for reading proficiency, student gender, and socioeconomic status, is there a significant difference in student performance on the end-of -course assessment in biology associated with race/ethnicity?

It was hypothesized that there will be a significant difference between the race/ethnicity of a student and their performance on the end-of -course assessment in biology. It is hypothesized that students from Caucasian backgrounds will score significantly higher than those from other ethnic backgrounds. It is also hypothesized that students from Caucasian backgrounds will score significantly higher than those from other ethnic backgrounds after adjusting for reading proficiency, gender, and socioeconomic background.

Mean Biology I EOC T-scale scores for 105 White, 66 Black, 4 Hispanic, and 3 American Indian students were 53.77 (SD = 10.88), 42.23 (SD = 10.53), 44.75 (SD = 10.15), and 48 (SD = 8.66), respectively. Table 4-11 provides descriptive statistics on the influence of ethnicity on Biology I EOC T-scale scores. With alpha equal to 0.05, a one factor between subjects ANOVA indicated a significant effect of ethnicity and Biology EOC T-scale scores, F (3, 174) = 15.92, $p < 0.05$. Table 4-12 provides the results of the one-way ANOVA on ethnicity and Biology I EOC T-scale scores. Tukey post-hoc comparisons of the four groups (White, Black, Hispanic, and American Indian or Alaskan native) indicate that the White students (M = 53.77, 95% CI [51.67, 55.88]) scored significantly higher on the Biology I EOC than Black students (M = 42.23, 95% CI [39.64, 44.82]), $p < 0.05$. Comparisons between the Hispanic (M = 44.75, 95% CI

[28.61, 60.89]) and the other three groups were not statistically significant, $p > 0.05$. Comparisons between the American Indian or Alaskan native ($M = 48$, 95% CI [26.49, 69.51]) and the other three groups were not statistically significant, $p > 0.05$. Table 4-13 provides data on the Tukey post hoc analysis of the influence of ethnicity on Biology I EOC T-scale scores.

The analysis of covariance (ANCOVA) was used to investigate the hypothesis that the observed difference in mean scores of the Biology I EOC T-scale scores is influenced by the differences in ethnicity, with the covariates, reading proficiency, socioeconomic status, and gender. The ANCOVA, however, found that the means for students of varying ethnicities was not significant, ($F = 0.63$, $p > 0.05$) - that is, after the effect of reading proficiency, socioeconomic status, and gender has been accounted for. Based upon the findings; hypothesis two was rejected.

- **Research Question 3**. Is there a significant difference in student performance on the end-of-course assessment in biology associated with socioeconomic status? After adjusting for reading proficiency, race/ethnicity, and student gender, is there a significant difference in student performance on the end-of -course assessment in biology associated with socioeconomic status?

It was hypothesized that there will be a significant difference in student performance on the end-of -course assessment in biology associated with socioeconomic status. It is hypothesized that students from higher socioeconomic backgrounds will score significantly higher than those from lower socioeconomic backgrounds. It is also hypothesized that students from higher socioeconomic backgrounds will score significantly higher than those from lower socioeconomic backgrounds after adjusting for student gender, race/ethnicity, and reading proficiency.

Mean Biology I EOC T-scale scores for students eligible for free lunch, reduced-price lunch, and not eligible for free/reduced-price lunch were 42.57 ($SD = 10.99$) for 69

students, 48.59 (*SD* = 9.31) for 17 students, and 54.27 (*SD* = 10.72) for 92 students, respectively. Table 4-14 provides additional descriptive statistics on the influence of socioeconomic status on Biology I EOC T-scale scores. With alpha equal to 0.05, a one factor between subjects ANOVA indicated a significant effect of free and reduced-price lunch eligibility and Biology EOC T-scale scores, F (2, 175) = 23.60, MSE = 114.65, $p < 0.05$. Table 4-15 provides further data on the one-way ANOVA results on socioeconomic status and Biology I EOC T-scale scores. Tukey post-hoc comparisons of the three groups (free, reduced, and not eligible) indicate that the students not eligible for free or reduced-price lunch (M = 54.27, 95% CI [52.05, 56.49]) scored significantly higher on the Biology I EOC than students who were eligible for free lunch (M = 42.57, 95% CI [39.92, 45.21]), $p < 0.05$. Comparisons between the reduced-price lunch (M = 48.59, 95% CI [43.80, 53.37]) and the other two groups were not statistically significant when alpha was set at 0.05 ($p > 0.05$), although students eligible for reduced-price lunch scored significantly higher than those eligible for free lunch at the $p < 0.10$ level. Table 4-16 provides additional data on a Tukey post hoc analysis of the influence of socioeconomic status on Biology I EOC T-scale scores.

The analysis of covariance (ANCOVA) was used to investigate the hypothesis that the observed difference in mean scores of the Biology I EOC T-scale scores is influenced by the differences in socioeconomic status, with the covariates, reading proficiency, ethnicity, and gender. The ANCOVA, however, found that the marginal means for students eligible for free and reduced-price lunch and those ineligible are not significant, (F = 1.84, $p > 0.05$) - that is, after the effect of reading proficiency, ethnicity,

and gender has been accounted for. Based upon the findings; hypothesis three was rejected.

- **Research Question 4.** Is there a significant difference in student performance on the end-of-course assessment in biology associated with gender? After adjusting for reading proficiency, race/ethnicity, and socioeconomic status, is there a significant difference in student performance on the end-of -course assessment in biology associated with gender?

It was hypothesized that there will be a significant difference in student performance on the end-of -course assessment in biology associated with gender. It is hypothesized that males will score significantly higher than females on the Biology I end-of -course assessment. It is hypothesized that males will score significantly higher on the end-of -course assessment in biology after adjusting for reading proficiency, race/ethnicity, and socioeconomic background.

Mean Biology I EOC T-scale scores was 50.68 (SD = 11.53) for 96 females and 47.45 (SD = 12.36) for 82 males. Table 4-17 provides additional data regarding descriptive statistics on the influence of gender on Biology I EOC T-scale scores. With alpha equal to 0.05, a one tailed independent samples t-test indicated no significant effect of gender on Biology I EOC T-scale scores, t (176) = -1.80, $p > 0.05$. Levene's Test for Equality of Variances indicates variances for males and females do not differ significantly from each other (p = 0.63). Table 4-18 provides additional data on the independent samples t-test to determine the influence of gender on Biology I EOC T-scale scores.

The analysis of covariance (ANCOVA) was used to investigate the hypothesis that the observed difference in mean scores of the Biology I EOC T-scale scores is influenced by gender differences and the covariates, reading proficiency, ethnicity, and socioeconomic status. The ANCOVA indicated that the marginal means for both male

and female students are not significant, ($F = 1.93$, $p > 0.05$) - that is, after the effect of reading proficiency, ethnicity, and socioeconomic status had been accounted for. Based upon the findings; hypothesis four was rejected.

Research Question 5. What are the best predictors of student performance on the Biology end-of-course assessment?

It was hypothesized that reading proficiency, ethnicity, and socioeconomic status will serve as significant predictors of student performance on the Biology end-of-course assessments.

During the stepwise regression procedure, a correlation matrix was produced that indicates the relationship between each variable and the dependent variable (see Table 4-19). Statistically significant correlations between Biology T-scale scores and the following variables were found: FCAT Reading developmental scale scores (0.73), student gender (0.13), White students (0.46), Black students (-0.45), students not eligible for free and reduced-price lunch (0.44), students eligible for free lunch (-0.44), minority students (0.46), and students enrolled in the 9th and 10th grade (-0.29), and the 11th grade (-0.16), $p < 0.05$. Other notable significant correlations existed between FCAT reading developmental scale scores and grade level (-0.24), White students (0.49), Black students (-0.50), students not eligible for free or reduced-price lunch (0.42), students eligible for free lunch (-0.42), minority students (0.49), and students enrolled in the 9th and 10th grade (-0.15), and the 11th grade (-0.15), $p < 0.05$.

Table 4- 20 reports the summary results of the four models as an outcome of the stepwise multiple regression procedure. As the number of predictors increased, the R value also increased from 0.73 to 0.77. Therefore, model 4 (which includes the

predictive variables FCAT reading developmental scale scores, grade levels 9-11, and students eligible for free lunch) had a statistically significant correlation of 0.77. The *R* square for this model was 0.60, therefore indicating that 60% of the variance in the Biology T-scale scores can be explained by the four predictive variables (FCAT reading developmental scale scores, grade levels 9-11, and students eligible for free lunch). The adjusted *R* square value (which takes into consideration the sample size and number of predictors) remained at 0.59, thus illustrating that 59% of the variance in biology scores is still attributable to the four aforementioned variables. Other analysis using an ANOVA indicated a significant effect of FCAT Reading developmental scale scores, grade level, and students eligible for free lunch and Biology EOC T-scale scores, F (4, 173) = 63.58, $p < 0.05$ (see Table 4-21).

After adding and removing other predictive variables that were statistically significant on Biology T-scale scores in the previous correlation matrix (from Table 4-19) such as White students (RegressionWhite), Black students (RegressionBlack), students not eligible for free and reduced-price lunch (RegressionNOTeligible), and minority students (Regressionminority); the researcher selected the model that had the largest *R* squared and adjusted *R*-squared values, which was statistically significant, and removed highly correlated predictors. Notable significant correlations that affected the selection of variables for the model included: the high correlation between students not eligible for free and reduced-price lunch and White students (0.50), Black students and eligibility for free lunch (0.42), White students and FCAT Reading developmental scale scores (0.49), and Black students and FCAT Reading developmental scale scores (-0.50); thereby indicating that there are intercorrelations or multicollinearity amongst the

predictor variables. The researcher, in order to eliminate redundancy of predictor variables or multicollinearity as well as achieve parsimony of the regression model, removed the following predictive variables from the model- White students, Black students, and students not eligible for free and reduced-price lunch.

Consequently, four predictive variables remained in the model, notably FCAT Reading developmental scale scores, grade levels 9 and 10, and grade level 11, and students eligible for free lunch. Table 4-22 summarizes the relationship of each of these variables to each other and the dependent variable, Biology T-scale scores, all of which were significant at the 0.05 level. According to the correlation matrix, Biology T-scale scores were significantly related to FCAT Reading developmental scale scores (0.73), eligibility for free lunch (-0.44), students in the 9th and 10th grade (-0.29), and students in the 11th grade (-0.16). Therefore indicating that students with higher FCAT Reading developmental scale scores had higher Biology T-scale scores. Furthermore, students eligible for free lunch and those in higher grades earned lower Biology T-scale scores. Based upon the findings; hypothesis five was rejected.

A summary of the model (Table 4-23) indicates that all four predictive variables collectively contribute to 60% of the variance in Biology T-scale scores. The adjusted *R* squared value is also 59%, after factoring in the sample size and number of predictive variables. There is also a significant correlation between the four predictor variables and the dependent variable ($R = 0.77$). Last, the Durbin-Watson test is a test that is used to determine whether the residuals from a multiple regression are independent. The statistic of 1.52 falls within the critical values of the Durbin-Watson statistic based on

sample size and number of regressors, and therefore the residuals are independent and there are no meaningful serial correlations.

The results of an ANOVA (Table 4-24) show that with an alpha equal to 0.05, a significant effect of FCAT Reading developmental scale scores, grade level, and eligibility for free lunch and Biology EOC T-scale scores, F (4, 177) = 63.58, MSE = 3790.25, $p < 0.05$. Table 4-25 provides additional information regarding the contribution of each predictor variable to the model. According to the results using standardized beta (β) coefficients, FCAT Reading developmental scale scores contribute 0.63 to the predictive model, whereas students eligible for free lunch -0.13, students enrolled in the 9th and 10th grade -0.20, and students enrolled in the 11th grade -0.11, $p < 0.05$. Therefore, students with higher FCAT Reading developmental scale scores had higher Biology T-scale scores, after controlling for the other variables in the model. The eligibility for free lunch and grade level had a significant negative weight, indicating that students who were eligible for free lunch and were in higher grades had lower Biology T-scale scores (a suppressor effect). Unstandardized B coefficients also indicate that FCAT Reading developmental scale scores contributes 0.38 to the predictive model, students eligible for free lunch -3.09, students enrolled in 9th and 10th grade -5.36, and students enrolled in the 11th grade -8.01. Therefore, for each added point or every unit increase on the FCAT Reading developmental scale score, Biology T-scale scores will increase by 0.38 (holding all other variables constant). Conversely, eligibility for free lunch, enrollment in the 10th and 11th grade reduces biology scores by 3.09, 5.36, and 8.01 points respectively (holding all other variables constant).

To ensure that all the assumptions of a multiple regression were met- that is, that variables are normally distributed, an assumption of a linear relationship between the independent and dependent variable, the measurement of variables without error, and the assumption of homoscedasticity; a combination of various visual aids were created. To test the normal distribution of the variables, a histogram of the frequency of regression standardized residuals was created (Figure 4-2). The visual inspection indicates that the distribution is indeed relatively normal. To test whether a linear relationship between the independent and dependent variables existed, a normal P-P plot was used, which indicates a linear relationship (Figure 4-3). To test whether the variables were measured without error, a reduction in the number of independent variables in the model was used as well as the inclusion of the adjusted R^2. To test the assumption of homoscedasticity, a visual inspection of a plot of the standardized residuals by the regression standardized predicted value was undertaken (Figure 4-4) which indicates a relatively even distribution, thus satisfying this assumption. Collectively, all assumptions were met, thus reducing the probability of Type I and II errors.

The regression equation therefore is:

Biology score = (0.379*FCAT reading developmental scale score) + (-5.356* student enrolled in the 9th or 10th grade) + (-8.009 * student enrolled in the 11th grade) + (-3.085* student eligible for free lunch) - 36.697

Summary

The purpose of this section is to describe the quantitative results of the data analysis of this study. The findings indicated that FCAT Reading developmental scale scores are strongly correlated with Biology I EOC T-scale scores. Students who

attained high FCAT Reading developmental scale scores scored significantly higher on the Biology I EOC assessment than students who attained low FCAT Reading developmental scale scores. However, there were no significant differences found on the influence of ethnicity on Biology I EOC T-scale scores when scores were adjusted for reading proficiency, socioeconomic status, and gender. No significant differences were also found on the influence of socioeconomic status on Biology I EOC T-scale scores when scores were adjusted for reading proficiency, ethnicity, and gender. Last, no significant differences were found on the influence of gender on Biology I EOC T-scale scores when scores were adjusted for reading proficiency, ethnicity, and socioeconomic status.

The significant predictors for biology scores on the end-of-course assessment included: FCAT Reading developmental scale scores, grade level, and eligibility for free lunch. FCAT Reading developmental scale scores, grade level, and eligibility for free lunch explained 60% of the variance in biology scores. FCAT Reading developmental scale scores, grade level, and eligibility for free lunch collectively, is also significantly correlated to biology scores. FCAT Reading developmental scale scores was the most significant contributor to the predictive model, followed by enrollment in either the 9th or 10th grade, eligibility for free lunch, and enrollment in the 11th grade.

Students with higher FCAT Reading developmental scale scores are predicted to earn higher biology scores whilst students enrolled in higher grades and eligible for free lunch are predicted to earn lower biology scores, after controlling for other variables in the study. Additionally, for every added point or every unit increase on the FCAT Reading developmental scale score, Biology T-scale scores will increase by 0.38

(holding all other variables constant). Conversely, eligibility for free lunch and enrollment in the 10th and 11th grade reduces biology scores by 3.09, 5.36, and 8.01 points, respectively (holding all other variables constant).The regression equation is: Biology score = (0.379*FCAT reading developmental scale score) + (-5.356* student enrolled in the 9th or 10th grade) + (-8.009 * student enrolled in the 11th grade) + (-3.085* student eligible for free lunch) - 36.697

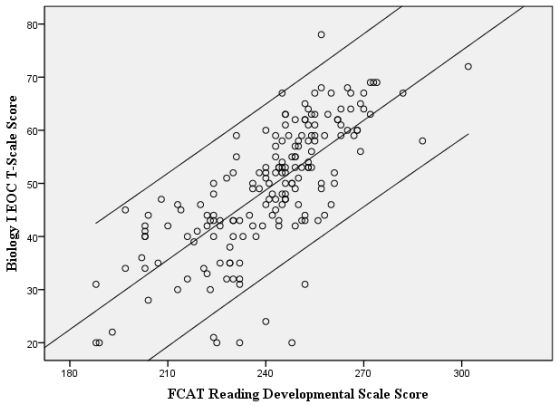

Figure 4-1. The relationship between FCAT Reading developmental scale scores and Biology I EOC T-scale scores

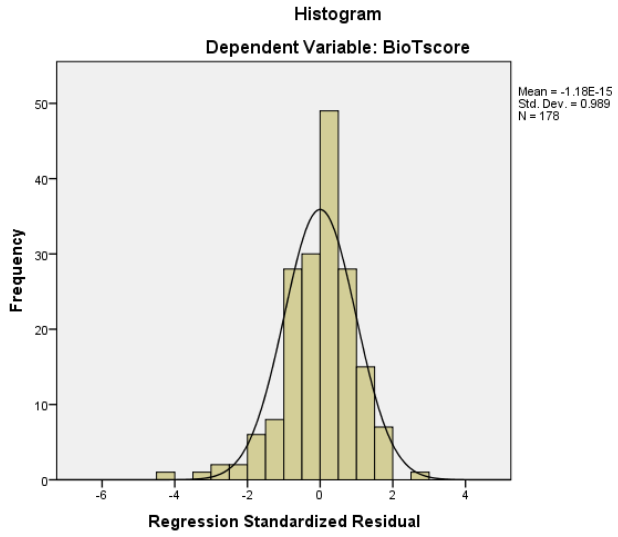

Figure 4-2. A histogram of the frequency of regression standardized residuals obtained from the initial stepwise regression analysis

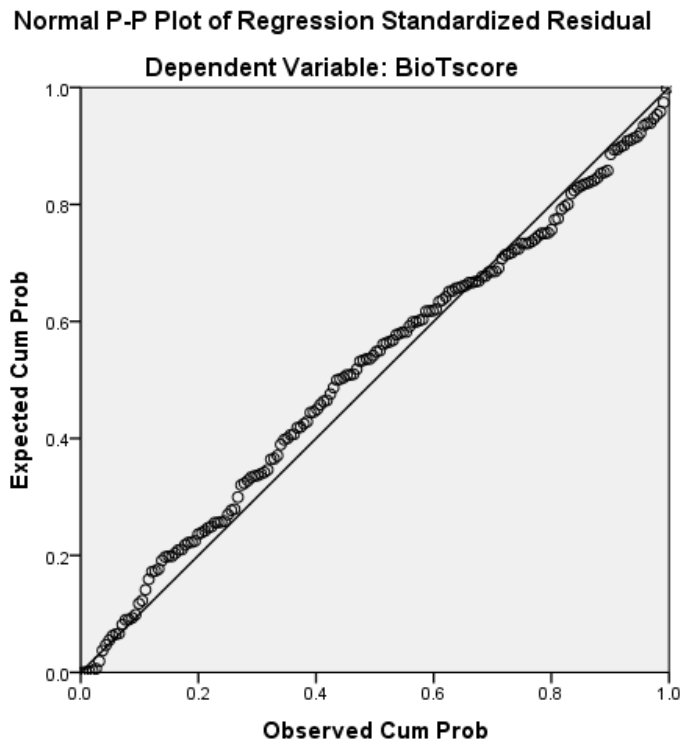

Figure 4-3. A Normal P-P plot of regression standardized residual that was used to determine how well a specific distribution fits the observed data

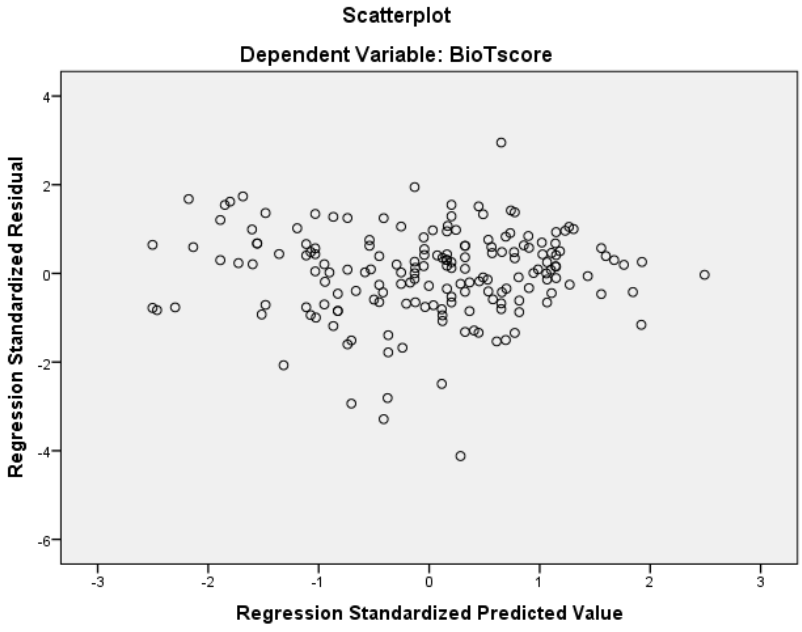

Figure 4-4. A scatterplot used to visually determine homoscedasticity of the data used in the initial stepwise regression analysis

Table 4-1. FCAT Reading level distribution by grade

		Grade Level			Total
		9	10	11	
FCAT Reading Level	1	2	34	2	38
	2	7	32	3	42
	3	15	39	0	54
	4	16	20	0	36
	5	3	5	0	8
Total		43	130	5	178

Table 4-2. Descriptive statistics on the influence of FCAT Reading level on Biology I EOC T-scale scores

Descriptives

Dependent variable: Biology T-scale score

FCAT Reading level	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	38	37.39	7.876	1.278	34.81	39.98	20	50
2	42	42.83	9.234	1.425	39.96	45.71	20	59
3	54	52.96	8.242	1.122	50.71	55.21	20	67
4	36	59.50	7.374	1.229	57.01	61.99	43	78
5	8	66.75	4.367	1.544	63.10	70.40	58	72
Total	178	49.19	11.997	.899	47.42	50.97	20	78

Table 4-3. One-way ANOVA results on FCAT reading levels and Biology I EOC T-scale scores

ANOVA

Biology T-scale scores

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	14046.167	4	3511.542	53.162	.000
Within Groups	11427.338	173	66.054		
Total	25473.506	177			

Table 4-4. Tukey post hoc analysis of the influence of FCAT Reading developmental scale scores on Biology I EOC T-scale scores

Multiple Comparisons

Dependent Variable: Biology T-scale score

Tukey HSD

Table 4-4. Continued.

(I) FCAT Reading Level	(J) FCAT Reading Level	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
1	2	-5.439*	1.820	.026	-10.45	-.42
	3	-15.568*	1.721	.000	-20.31	-10.82
	4	-22.105*	1.890	.000	-27.32	-16.89
	5	-29.355*	3.161	.000	-38.07	-20.64
2	1	5.439*	1.820	.026	.42	10.45
	3	-10.130*	1.672	.000	-14.74	-5.52
	4	-16.667*	1.846	.000	-21.76	-11.58
	5	-23.917*	3.135	.000	-32.56	-15.27
3	1	15.568*	1.721	.000	10.82	20.31
	2	10.130*	1.672	.000	5.52	14.74
	4	-6.537*	1.749	.002	-11.36	-1.72
	5	-13.787*	3.079	.000	-22.27	-5.30
4	1	22.105*	1.890	.000	16.89	27.32
	2	16.667*	1.846	.000	11.58	21.76
	3	6.537*	1.749	.002	1.72	11.36
	5	-7.250	3.177	.156	-16.01	1.51
5	1	29.355*	3.161	.000	20.64	38.07
	2	23.917*	3.135	.000	15.27	32.56
	3	13.787*	3.079	.000	5.30	22.27
	4	7.250	3.177	.156	-1.51	16.01

*. The mean difference is significant at the 0.05 level.

Table 4-5. ANCOVA results for FCAT Reading level and Biology I EOC T-scale scores

Tests of Between-Subjects Effects

Dependent Variable: Biology T-scale score

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	14538.355[a]	7	2076.908	32.288	.000	.571
Intercept	14911.548	1	14911.548	231.818	.000	.577
Gender	112.080	1	112.080	1.742	.189	.010
Ethnicity	89.755	1	89.755	1.395	.239	.008
Free and Reduced Lunch Eligibility	196.811	1	196.811	3.060	.082	.018
FCAT Reading Level	7686.254	4	1921.564	29.873	.000	.413
Error	10935.150	170	64.324			
Total	456190.000	178				
Corrected Total	25473.506	177				

a. R Squared = .571 (Adjusted R Squared = .553)

Table 4-6. Tukey post hoc analysis of the influence of ethnicity on FCAT Reading developmental scale scores

Multiple Comparisons

Dependent Variable: FCAT Reading Developmental Scale Score

Tukey HSD

(I) Ethnicity	(J) Ethnicity	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
White	Black	21.126*	2.751	.000	13.99	28.26
	Hispanic	11.921	8.922	.541	-11.22	35.06
	American Indian or Alaskan native	3.838	10.255	.982	-22.76	30.44
Black	White	-21.126*	2.751	.000	-28.26	-13.99
	Hispanic	-9.205	9.018	.738	-32.60	14.19
	American Indian or Alaskan native	-17.288	10.338	.342	-44.11	9.53
Hispanic	White	-11.921	8.922	.541	-35.06	11.22
	Black	9.205	9.018	.738	-14.19	32.60
	American Indian or Alaskan native	-8.083	13.376	.931	-42.78	26.61
American Indian or Alaskan native	White	-3.838	10.255	.982	-30.44	22.76
	Black	17.288	10.338	.342	-9.53	44.11
	Hispanic	8.083	13.376	.931	-26.61	42.78

*. The mean difference is significant at the 0.05 level.

Table 4-7. ANCOVA results for ethnicity and FCAT Reading developmental scale scores

Tests of Between-Subjects Effects

Dependent Variable: FCAT Reading Developmental Scale Score

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	21870.421[a]	5	4374.084	15.139	.000	.306
Intercept	404168.234	1	404168.234	1398.828	.000	.891
Gender	34.044	1	34.044	.118	.732	.001
Free and Reduced Lunch Eligibility	3603.971	1	3603.971	12.473	.001	.068
Ethnicity	7800.229	3	2600.076	8.999	.000	.136
Error	49696.574	172	288.934			
Total	10410467.000	178				
Corrected Total	71566.994	177				

a. R Squared = .306 (Adjusted R Squared = .285)

Table 4-8. Tukey post hoc analysis of the influence of socioeconomic status on FCAT Reading developmental scale scores

Multiple Comparisons

Dependent Variable: FCAT Reading Developmental Scale Score

Tukey HSD

(I) What is the eligibility of the student regarding free/reduced lunch?	(J) What is the eligibility of the student regarding free/reduced lunch?	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Free	Reduced	-9.623	4.914	.126	-21.24	1.99
	Not eligible	-18.786*	2.890	.000	-25.62	-11.95
Reduced	Free	9.623	4.914	.126	-1.99	21.24
	Not eligible	-9.163	4.791	.138	-20.49	2.16
Not eligible	Free	18.786*	2.890	.000	11.95	25.62
	Reduced	9.163	4.791	.138	-2.16	20.49

*. The mean difference is significant at the 0.05 level.

Table 4-9. ANCOVA results for socioeconomic status and FCAT Reading developmental scale scores

Tests of Between-Subjects Effects

Dependent Variable: FCAT Reading Developmental Scale Score

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	17768.985[a]	4	4442.246	14.285	.000	.248
Intercept	593592.022	1	593592.022	1908.833	.000	.917
Ethnicity	3697.918	1	3697.918	11.892	.001	.064
Gender	135.660	1	135.660	.436	.510	.003
Free and Reduced Lunch Eligibility	7260.215	2	3630.108	11.673	.000	.119
Error	53798.010	173	310.971			
Total	10410467.000	178				
Corrected Total	71566.994	177				

a. R Squared = .248 (Adjusted R Squared = .231)

Table 4-10. Independent samples *t*-test to determine the influence of gender on FCAT Reading developmental scale scores

Independent Samples Test

Table 4-10. Continued.

		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
FCAT Reading Developmental Score	Equal variances assumed	3.197	.075	-1.043	176	.298	3.023	-9.119	2.812
	Equal variances not assumed			-1.030	160.158	.304	3.061	-9.199	2.892

Table 4-11. Descriptive statistics on the influence of ethnicity on Biology I EOC T-scale scores

Descriptives

Biology T-scale scores

Ethnicity	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
White	105	53.77	10.883	1.062	51.67	55.88	20	78
Black	66	42.23	10.532	1.296	39.64	44.82	20	64
Hispanic	4	44.75	10.145	5.072	28.61	60.89	35	59
American Indian or Alaskan native	3	48.00	8.660	5.000	26.49	69.51	43	58
Total	178	49.19	11.997	.899	47.42	50.97	20	78

Table 4-12. One-way ANOVA results on ethnicity and Biology I EOC T-scale scores

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5486.650	3	1828.883	15.922	.000
Within Groups	19986.855	174	114.867		
Total	25473.506	177			

Table 4-13. Tukey post hoc analysis of the influence of ethnicity on Biology I EOC T-scale scores

Multiple Comparisons
Dependent Variable: Biology T-scale score
Tukey HSD

(I) Ethnicity	(J) Ethnicity	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
White	Black	11.544*	1.684	.000	7.18	15.91
	Hispanic	9.021	5.460	.352	-5.14	23.18
	American Indian or Alaskan native	5.771	6.276	.794	-10.51	22.05
Black	White	-11.544*	1.684	.000	-15.91	-7.18
	Hispanic	-2.523	5.519	.968	-16.84	11.79
	American Indian or Alaskan native	-5.773	6.327	.798	-22.18	10.64
Hispanic	White	-9.021	5.460	.352	-23.18	5.14
	Black	2.523	5.519	.968	-11.79	16.84
	American Indian or Alaskan native	-3.250	8.186	.979	-24.48	17.98
American Indian or Alaskan native	White	-5.771	6.276	.794	-22.05	10.51
	Black	5.773	6.327	.798	-10.64	22.18
	Hispanic	3.250	8.186	.979	-17.98	24.48

Table 4-14. Descriptive statistics on the influence of socioeconomic status on Biology I EOC T-scale scores

Biology T-scale score

Categories	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound	Minimum	Maximum
Free	69	42.57	10.994	1.324	39.92	45.21	20	63
Reduced	17	48.59	9.308	2.257	43.80	53.37	32	64
Not eligible	92	54.27	10.720	1.118	52.05	56.49	20	78
Total	178	49.19	11.997	.899	47.42	50.97	20	78

Table 4-15. One-way ANOVA results on socioeconomic status and Biology I EOC T-scale scores

ANOVA

Biology T-scale score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5410.225	2	2705.112	23.595	.000
Within Groups	20063.281	175	114.647		
Total	25473.506	177			

Table 4-16. Tukey post hoc analysis of the influence of socioeconomic status on Biology I EOC T-scale scores

Multiple Comparisons

Dependent Variable: Biology T-scale scores

Tukey HSD

(I) What is the eligibility of the student regarding free/reduced lunch?	(J) What is the eligibility of the student regarding free/reduced lunch?	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Free	Reduced	-6.023	2.899	.097	-12.88	.83
	Not eligible	-11.707*	1.705	.000	-15.74	-7.68
Reduced	Free	6.023	2.899	.097	-.83	12.88
	Not eligible	-5.684	2.827	.113	-12.37	1.00
Not eligible	Free	11.707*	1.705	.000	7.68	15.74
	Reduced	5.684	2.827	.113	-1.00	12.37

*. The mean difference is significant at the 0.05 level.

Table 4-17. Descriptive statistics on the influence of gender on Biology I T-scale scores

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Biology T-scale score	males	82	47.45	12.361	1.365
	females	96	50.68	11.533	1.177

Table 4-18. Independent samples *t*-test to determine the influence of gender on Biology I EOC T-scale scores

	t-test for Equality of Means						
	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						Lower	Upper

Table 4-18. Continued.

		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Biology T-scale score	Equal variances assumed	-1.799	176	.074	-3.226	1.793	-6.764	.312
	Equal variances not assumed	-1.790	167.354	.075	-3.226	1.802	-6.784	.333

Table 4-19. Correlation matrix indicating the relationship between possible predictive variables and Biology I EOC T-scale scores

		BioTscore	RDevScore	RegressWhite	RegressBlack	RegressHispanic	RegressAmInAlNa	RegressGENDER	RegressNOTeligible	Regressreducedeligible
Pearson Correlation	BioTscore	1.000	.733	.459	-.447	-.056	-.013	.134	.439	-.016
	RDevScore	.733	1.000	.488	-.496	-.028	.028	.078	.421	-.016
	RegressWhite	.459	.488	1.000	-.921	-.182	-.157	.077	.497	-.157
	RegressBlack	-.447	-.496	-.921	1.000	-.116	-.101	-.107	-.491	.146
	RegressHispanic	-.056	-.028	-.182	-.116	1.000	-.020	.064	-.005	-.049
	RegressAmInAlNa	-.013	.028	-.157	-.101	-.020	1.000	.033	-.048	.106
	RegressGENDER	.134	.078	.077	-.107	.064	.033	1.000	.076	-.006

Table 4-19. Continued.

	BioTscore	RDevScore	RegressWhite	RegressBlack	RegressHispanic	RegressAmInAlNa	RegressGENDER	RegressNOTeligible	Regressreducedeligible
RegressNOTeligible	.439	.421	.497	-.491	-.005	-.048	.076	1.000	-.336
Regressreducedeligible	-.016	-.016	-.157	.146	-.049	.106	-.006	-.336	1.000
Regressfreelunch	-.441	-.422	-.415	.416	.035	-.015	-.074	-.823	-.259
Regressminority	.459	.488	1.000	-.921	-.182	-.157	.077	.497	-.157
Regress10	-.287	-.148	-.121	.178	-.079	-.117	-.257	-.208	.025
Regress11	-.159	-.149	-.204	.221	-.026	-.022	-.048	-.176	.176
RDevScore	.000	.	.000	.000	.353	.354	.149	.000	.415
RegressWhite	.000	.000	.	.000	.008	.018	.153	.000	.018
RegressBlack	.000	.000	.000	.	.061	.091	.077	.000	.026
RegressHispanic	.228	.353	.008	.061	.	.396	.198	.473	.257
RegressAmInAlNa	.431	.354	.018	.091	.396	.	.329	.262	.080
RegressGENDER	.037	.149	.153	.077	.198	.329	.	.156	.466

Table 4-19. Continued.

	BioTscore	RDevScore	RegressWhite	RegressBlack	RegressHispanic	RegressAmInAlNa	RegressGENDER	RegressNOTeligible	Regressreducedeligible
RegressNOTeligible	.000	.000	.000	.000	.473	.262	.156	.	.000
Regressreducedeligible	.414	.415	.018	.026	.257	.080	.466	.000	.
Regressfreelunch	.000	.000	.000	.000	.322	.423	.162	.000	.000
Regressminority	.000	.000	.000	.000	.008	.018	.153	.000	.018
Regress10	.000	.024	.054	.009	.148	.060	.000	.003	.369
Regress11	.017	.023	.003	.001	.366	.384	.264	.009	.009

Table 4-20. Results of stepwise regression model indicating significant predictive variables on Biology I EOC T-scale scores

Model Summary[e]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.733[a]	.538	.535	8.179	
3	.764[c]	.584	.577	7.800	
4	.771[d]	.595	.586	7.721	1.513

a. Predictors: (Constant), RDevScore

c. Predictors: (Constant), RDevScore, Regress10, Regressfreelunch

d. Predictors: (Constant), RDevScore, Regress10, Regressfreelunch, Regress11

e. Dependent Variable: BioTscore

Table 4-21. ANOVA results of stepwise regression model indicating significant predictive variables on Biology I EOC T-scale scores

ANOVA[a]

Table 4-21. Continued.

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	13700.338	1	13700.338	204.810	.000[b]
	Residual	11773.168	176	66.893		
	Total	25473.506	177			
2	Regression	14524.794	2	7262.397	116.079	.000[c]
	Residual	10948.711	175	62.564		

Table 4-21. Continued.

	Total	25473.506	177			
3	Regression	14886.382	3	4962.127	81.553	.000[d]
	Residual	10587.124	174	60.846		
	Total	25473.506	177			
4	Regression	15161.011	4	3790.253	63.584	.000[e]
	Residual	10312.494	173	59.610		
	Total	25473.506	177			

a. Dependent Variable: BioTscore

b. Predictors: (Constant), RDevScore

c. Predictors: (Constant), RDevScore, Regress10

d. Predictors: (Constant), RDevScore, Regress10, Regressfreelunch

e. Predictors: (Constant), RDevScore, Regress10, Regressfreelunch, Regress11

Table 4-22. Correlation matrix results of regression model indicating significant predictive variables on Biology I EOC T-scale scores

Correlations

		BioTscore	RDevScore	Regressfreelunch	Regress10	Regress11
Pearson Correlation	BioTscore	1.000	.733	-.441	-.287	-.159
	RDevScore	.733	1.000	-.422	-.148	-.149
	Regressfreelunch	-.441	-.422	1.000	.198	.074
	Regress10	-.287	-.148	.198	1.000	-.280
	Regress11	-.159	-.149	.074	-.280	1.000
Sig. (1-tailed)	BioTscore	.	.000	.000	.000	.017
	RDevScore	.000	.	.000	.024	.023
	Regressfreelunch	.000	.000	.	.004	.163
	Regress10	.000	.024	.004	.	.000
	Regress11	.017	.023	.163	.000	.

Table 4-23. Results of regression model indicating significant predictive variables on Biology I EOC T-scale scores

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.771[a]	.595	.586	7.721	1.513

a. Predictors: (Constant), Regress11, Regressfreelunch, Regress10, RDevScore

b. Dependent Variable: BioTscore

Table 4-24. ANOVA results of regression model indicating significant predictive variables on Biology I EOC T-scale scores

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	15161.011	4	3790.253	63.584	.000[b]
	Residual	10312.494	173	59.610		
	Total	25473.506	177			

a. Dependent Variable: BioTscore

Table 4-24. Continued.

b. Predictors: (Constant), Regress11, Regressfreelunch, Regress10, RDevScore

Table 4-25. Results of regression model indicating significant predictive variables on Biology I EOC T-scale scores

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-36.697	8.219		-4.465	.000		
	RDevScore	.379	.032	.634	11.706	.000	.797	1.255
	Regressfreelunch	-3.085	1.328	-.126	-2.323	.021	.800	1.250
	Regress10	-5.356	1.405	-.199	-3.812	.000	.861	1.161
	Regress11	-8.009	3.731	-.111	-2.146	.033	.881	1.135

CHAPTER 5
DISCUSSION

Florida's Next Generation Strategic Plan operates under the premise that measuring student achievement through EOC assessments would serve as a tool of ensuring students are globally competitive for college and future careers (Florida Department of Education, 2005b). Starting in the 2012-2013 academic year, the state requires students to pass the one such assessment, the Biology EOC, with a satisfactory score or higher to obtain the course credit required to graduate (The Florida Legislature, 2012). Like similar assessments, the Biology EOC influences student outcomes such as retention and graduation rates--which are data indicators for school accountability that determines overall school grades (The Florida Legislature, 2012) and serve as the basis of teacher evaluations (The Florida Legislature, 2010).

This increased accountability makes the ability to make more informed decisions increasingly important. Thus, the purpose of this study is two-fold: (1) to examine four variables-- reading proficiency, ethnicity, socioeconomic status, and gender--to determine their influence on biology achievement as reflected on the results from the Florida Biology I EOC assessment at a Title 1 school, and (2) to predict which student characteristics are most influential on biology scores.

Of all four variables identified, reading proficiency (which is determined by FCAT Reading developmental scale scores) was the sole variable that had a significant influence on Biology I EOC T-scale scores. Results indicated that students scoring 3 or higher on FCAT Reading scored significantly higher on the Biology EOC than their peers who obtained lower FCAT reading scores. This finding regarding the importance of reading proficiency on science achievement is consistent with previous research that

more proficient readers obtained significantly higher scores on science assessments than their less proficient peers (Cromley, 2009; Dempster & Reddy, 2007; Haught & Wall, 2004; Medina & Mishra, 1994; O'Reilly & McNamara, 2007).

However, identifying the mechanism involved in reading that contributes to the disparity in scores is beyond the scope of this research. Nor can any conclusions be unequivocally made that improving FCAT Reading scores will increase Biology EOC scores. Because the causal-comparative approach of this study relies on determining the reason for differences in Biology I EOC scores after the fact, no cause-and-effect conclusions can be definitively drawn due to the lack of randomization, manipulation, and control.

Although reading proficiency, as indicated by FCAT Reading levels, was found to be the sole variable that influences student achievement in biology, results indicate that reading proficiency itself is influenced by ethnicity and socioeconomic status. In the sample group, 80% of African Americans qualified for free and reduced-price lunch—compared to 28% of the White students and 50% of Hispanic students. Thus the group demonstrated a correlation of ethnicity and socioeconomic status. Furthermore, the study revealed the high correlation between students not eligible for free and reduced-price lunch and White students, Black students and eligibility for free lunch, White students and FCAT Reading developmental scale scores, and Black students and FCAT Reading developmental scale scores. Thus the research revealed strong correlations between ethnicity, socioeconomic status, and reading performance.

Nevertheless, results indicated that White students scored significantly higher than Black students on reading assessments and that students ineligible for free and

reduced-price lunch scored significantly higher than eligible students . This finding is consistent with previous studies (College Board, 2011; National Center for Education Statistics, 2011; U.S. Department of Education, National Center for Education Statistics, 2012). Therefore, although ethnicity and socioeconomic status did not have a direct impact on biology achievement, those variables indirectly influence it by directly affecting reading proficiency. Again, no indisputable conclusions can be made regarding how or why those effects exist since such questions lie outside the scope of this research.

This research also revealed that collectively as a model, FCAT Reading developmental scale scores, grade level, and eligibility for free lunch were factors that could predict biology achievement and explain 60% of the variance in biology scores. However, when each variable except reading was examined independently, each failed to reach a significance of 0.05. Nevertheless, these results emphasize the need for effective teacher evaluation design that takes into account student demographic composition so as not to penalize teachers who serve a student population that has historically underperformed academically. Additionally, although extant literature fails to quantify the institutional effect on biology achievement, the significant variance in biology scores highlights the potential challenges as well as opportunities involved in improving biology achievement.

One challenge is to make substantial improvements in reading proficiency to minimize the negative effects of socioeconomic status. However, this is also an opportunity because improvement in reading proficiency levels will reduce the negative effects of some student characteristics and will result in improved biology achievement.

According to the model, reading proficiency accounts for 63% of the variation in biology scores, therefore, it is recommended that schools provide adequate and quality professional development in reading strategies to biology teachers. The recent implementation of Common Core standards that emphasize the development of student vocabulary and reading comprehension may serve as a catalyst in the process, especially as those standards will be tested in the future. Thus, schools will need to divert additional resources to ensure teachers are not only adequately prepared for biology instruction, but also in incorporating reading strategies within the classroom.

Biology achievement was significantly correlated with ethnicity, but ethnicity was also significantly correlated with eligibility for free lunch. The inclusion of ethnicity in the model resulted in statistically insignificant values particularly compared to the greater statistical strength of socioeconomic status and of free lunch as the best predictor of socioeconomic status. The choice to include grade level in the model may seem unexpected; however, study results indicated that as grade level increased, reading proficiency declined, there were less White and students ineligible for free and reduced-price lunch, and the sample of Black students increased. Therefore grade level represents a conglomeration of the variables associated with ethnicity, reading proficiency, and socioeconomic status. Consequently, the results provide evidence supporting the significant impact of socioeconomic status on overall biology and reading achievement.

Another important finding in this study is the strong, positive, linear relationship between FCAT Reading scores and Biology I EOC scores. As FCAT Reading scores increased, Biology I EOC scores also increased. In fact, FCAT Reading developmental

scale scores served as the most significant predictor of biology achievement in the regression analysis. Although correlations refrain from making causal statements, they do indicate the presence of a relationship between the two variables. These relationships may be mediated by no, a sole, or a combination of covariates which serve to influence either variable. It is uncertain whether high FCAT Reading developmental scale scores results in higher Biology I EOC T-scale scores or if students with higher Biology I EOC T-scale scores perform better in reading assessments than their peers with lower Biology I EOC T-scale scores.

These results beg the question: Does increased reading proficiency help to accurately predict biology proficiency? Considering that the Biology I EOC test items are written at a 9th grade reading level (Florida Department of Education, 2011c), a certain level of reading proficiency is required. Furthermore, it is unclear whether the Biology EOC assessment is comprised of compound sentences and complex vocabulary that collectively increase sentence complexity. Consequently, in an attempt to measure biology proficiency, the test may inadvertently be assessing reading proficiency. In that case, the EOC would fail to provide a valid and fair assessment of the biology knowledge and skills of less proficient readers.

Until the test developers address this issue, this researcher proposes that increased reading proficiency may increase the probability that scores obtained are representative of student biology proficiency. However, it is possible that increased reading proficiency may improve scores beyond actual biology proficiency through the use of covariates such as background knowledge, ability to use context clues, and enhanced vocabulary. Thus test creators must ensure that the Biology EOC assesses

student knowledge of biology concepts and skills rather than unduly assessing reading proficiency. This research would suggest that test creators should use simple sentence structures with basic domain general vocabulary.

The nature of the Biology EOC sentence complexity is, however, beyond the scope of this study. On the other hand, educators in Florida can make use of the familiarity of the FCAT Reading assessment and resulting score interpretations indicated by the varying achievement levels identified in the study. Since FCAT Reading data is more easily accessible to teachers and administrators that Biology EOC assessment data, educators can use reading scores make better instructional decisions in biology to improve student comprehension and achievement. These results also provide an initial basis for identifying students at risk of failing the Biology I EOC assessment who, according to the data, are struggling readers with low FCAT Reading developmental scale scores. In addition, instructional activities that improve FCAT Reading developmental scale scores could be implemented in the classroom to potentially improve biology achievement.

Based on these results and previous studies, three possible explanations exist regarding the relationship between reading and biology proficiency: 1) reading proficiency results in biology proficiency, 2) biology proficiency results in reading proficiency, or 3) covariate(s) result in both reading proficiency and biology proficiency. An unequivocal conclusion to this question is beyond the scope of this study, but extant research may suggest the influence of covariates (Cromley, 2009). According to Cromley (2009), the high correlations between reading comprehension and science proficiency may be the result of plausible third factors that were not investigated in their

study, such as artifacts of measurement, instructional practices, the familiarity of students with standardized tests, or the alignment of home and school cultures or practices.

Cromley (2009) proposed additional covariates that may include the influence of background knowledge, reading comprehension strategies, general vocabulary, inference, and other products of extensive reading experience which increase science proficiency. This hypothesis is supported by recent research which indicates that specific instruction in reading strategies, vocabulary, and background knowledge have been shown to increase student achievement in science (Fang & Wei, 2010; Guthrie, Wigfield, Barbosa, Perencevich, Taboada, Davis, Scafiddi, &Tonks, 2004; Reaves, 2000; Shymansky, Yore, & Anderson, 2004; Vitale & Romance, 2012; Yore, Bisanz, & Hand, 2003). Cromley, Snyder-Hogan, and Luciw-Dubas (2010) also found that background knowledge, inference, and vocabulary had significant effects on biology achievement for high school students.

Additionally, students with more scientific background knowledge exhibit better reading comprehension of scientific texts due to the direct effect of background knowledge and the indirect effects of strategies and inference (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010). Consequently, according to Cromley (2009), by increasing reading comprehension of scientific texts, the degree of student proficiency in science increases (Cromley, 2009). However, the assessments of reading and science used in the Cromley (2009) study were heavily weighted toward primarily evaluating students utilizing inferences. Cromley, Snyder-Hogan, and Luciw-Dubas's 2010 study also utilized biology texts that relied heavily on inferences. It is uncertain whether a similar

heavy reliance on inferences is utilized in the Biology I EOC, so these results may not be as applicable. Thus the role of inference as a covariate influencing science or reading proficiency in this study is uncertain.

Implications

Findings of this study indicate that by improving reading proficiency, *all* students, regardless of ethnicity, socioeconomic status, or gender, will improve their achievement in biology. Therefore the results can assist educators and policymakers in high-stakes testing educational climates to integrate reading into the content areas at the secondary level. This would require a shift in the allocation of resources and would impact instructional practices. Such a shift could pose a challenge due to limited instructional time, potential inadequate teacher knowledge of reading pedagogy and its incorporation into the science classroom, and perceived threats to teacher autonomy.

However, the recent required incorporation of Common Core standards in Florida classrooms may serve as a catalyst to build teacher proficiency in reading instruction and students' acquisition of vocabulary. Considering that all teachers will be required to incorporate grade level specific reading strategies to satisfy Common Core requirements, only time will reveal whether students will be more prepared to handle the levels of reading proficiency required for the biology EOC. The implementation of Common Core standards also raises questions whether the biology EOC will become increasingly more challenging as a means of measuring student proficiency with the standards.

Regardless, since student performance serves as a significant influence on teacher evaluations, biology teachers need to become familiar with and incorporate reading strategies within the classroom so that student achievement is improved. In

fact, the study indicates that for every unit increase on the FCAT Reading developmental scale score, biology scores would increase by 0.38. Therefore, professional development in incorporating reading strategies is highly indicated, especially for teachers working in Title 1 schools. Of course, solely emphasizing reading to the neglect of biology content is not advisable. In fact, 39 students (22%) in the study who were proficient in reading failed to score in the top third of the biology assessment, so high quality biology content instruction is clearly essential. However, it is recommended that teachers infuse reading instruction into biology content to improve conceptual understanding and student achievement.

Test creators also need to be cognizant of the reading levels of the questions so that they accurately measuring biology achievement instead of both biology and reading proficiency. Failure to account for the assessment's increased reading demands t, may serve to reduce the numbers of African American and low socioeconomic students who graduate from high school and thus limiting their access to beneficial post-graduation opportunities such as higher-paying jobs and advanced education.

Educational policy runs the risk that the goal of creating high school graduates who are more globally competitive in careers and at the collegiate level has increasingly subjugated a subset of the student population (Florida Department of Education, 2005b). While legislative directives are aimed at increasing academic achievement of all students, further analysis must be conducted to determine whether this goal is being met and whether public policy is providing the resources needed to ensure that all students have access to resources as well as highly qualified and effective biology teachers.

For school administrators and educational policymakers, the implementation of increased stakes in successive administrations of the Biology end-of-course assessment may be accompanied by potential drops in school accountability indicators due to the performance of struggling readers unless institutional adjustments are made in terms of the allocation of time, resources, and personnel. Thus, it is essential that instructional leaders and administrators utilize data-driven decision making to analyze both the reading proficiency levels of both the incoming freshman class and all those students currently enrolled in Biology. This analysis is particularly important for Title 1 schools that serve populations that have historically underperformed in reading. Due to the study's sample size and methodology, it is unclear whether similar achievement gaps based on gender, race/ethnicity, and socioeconomic status would be found at other schools, particularly Title 1 schools. It is therefore also recommended that all schools analyze student achievement data to determine whether *all* students are being equally served by the instructional personnel and the biology curriculum. Institutional changes may include providing professional development in content area reading, culturally appropriate pedagogical practices, gender-neutral classrooms, and making instructional personnel changes, etc.

The results of the study also highlight potential issues for instructional evaluations and the recruitment and retention of biology teachers in Florida particularly in high poverty schools. According to Council of Chief State School Officers (2011), one potential reason for the implementation of end-of-course assessments is to determine teacher effectiveness. However, such a policy is fraught with numerous challenges. For example, this study and countless others have consistently demonstrated the

underperformance of low-income and minority students in science (ACT, 2010; Bankston & Caldas, 1998; Florida Department of Education, 2012b; Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008; Maerten-Rivera, Myers, Lee, & Penfield, 2010; National Center for Education Statistics, 2012; Rathbun & West, 2004; Riegle-Crumb, Moore, & Ramos-Wada, 2010; Strand, 2012). Recent changes in Florida's instructional evaluations make student achievement scores account for 50% of a teacher's evaluation, so that some teachers could be unfairly penalized for the demographic composition of the student body and its associated challenges. This is particularly critical as all new teachers hired after July 1, 2011 are offered annual contracts in which continued employment and compensation are tied to student outcomes (Harrison & Cohen-Vogel, 2012).

For this reason any use of test data to determine teacher effectiveness should control for prior student performance (Council of Chief State School Officers, 2011). However, the current teacher evaluation tool does not factor in student gains for biology teachers. Therefore a single test, the end-of-course assessment, will have a significant impact on a teacher's evaluation, continued employment, and opportunities for financial incentives. Research has indicated that an accumulation of classroom and school effects impact on student achievement, therefore relying on such a measure could produce a highly biased estimate of teacher effects (Sykes, Schneider, Plank, 2009). Additionally, decisions that impact on student achievement are made at various levels, such as federal, state, district, school, and classroom. For example, curriculum decisions may be made at the state and district level as well as other decisions such as

class size, teaching personnel, and budgets; yet such interactions are not considered in a teacher's evaluation (Sykes, Schneider, Plank, 2009).

According to Council of Chief State School Officers (2011), any teacher evaluation measurement should reflect the progress and growth of a student during the period of time the teacher provided instruction. This requires that two elements must be in place: availability of one or more prior scores and a suitable analytic method. However, even when a trustworthy measure of student progress is developed, other challenges have to be addressed in order to associate such results with teacher effectiveness, such as extraneous factors that threaten the interpretation that the teacher's behavior led to the observed gains, such as truancy, or family crisis during the instructional term (Council of Chief State School Officers, 2011).

Student characteristics account for the majority of the influence on biology achievement at the Title 1 school in this study, but teachers in this environment are ultimately held accountable for student achievement. Considering that research has documented the shortage of qualified instructional personnel in science in rural and Title 1 schools (Baker & Cooper, 2005; Murphy, DeArmond & Guin, 2003; National Science Foundation, 2012a), as well as their high attrition levels (Barrera, Braley & Slate, 2010) compared to their more affluent counterparts, policymakers need to devise a more suitable evaluation tool for science teachers working at Title 1 schools. Furthermore, policymakers must also ensure that schools serving populations of struggling students also receive adequate resources and effective professional development to address student deficiencies in reading. Policymakers must also ensure that they do not pose

additional burdens on such schools to attract and retain highly qualified instructional personnel.

One proposed solution to the discrepancy in measuring teacher effectiveness is the use of the Value-Added Modeling (VAM). This process utilizes several statistical techniques that seeks to link or establish causality between student test scores on a standardized test and a teacher's performance (Hill, Kapitula, & Umland, 2011). According to Florida Department of Education (2012q), the VAM model takes into account other factors that may affect the learning process and does not evaluate teachers based on student data derived from a single year or on a comparison of growth from one year to the next. Instead, teachers are evaluated based on the difference between the predicted performance and the actual performance of students. Thus, the model purports to "level the playing field" by statistically accounting for the differences in both the level of student proficiency and student characteristics that are assigned to teachers (Florida Department of Education, 2012q). Therefore, no teacher is unfairly advantaged or disadvantaged simply based on student composition.

However, Green III, Baker, & Oluwole (2012) cited multiple problems with the VAM approach. They noted the "instability of teacher ratings, classification and model prediction error, unreliable results from different 'standardized' tests, difficulties in isolating a single teacher's contribution to students' learning, the non-random assignment of students across teachers, schools, and districts, and the struggle for teachers to even receive VAM ratings" (Green III, Baker, & Oluwole, 2012, p. 6). Hill, Kapitula, & Umland's 2011 study also revealed that although VAM scores converged with the ratings from experts regarding their instruction, data also revealed that the

scores correlated with the student composition of a teacher's classroom. Thus the authors advocate the use of VAM scores alongside high-quality, discriminating observational systems or as a trigger for the use of observations, as VAM scores alone are not sufficient to identify teachers in need of remediation, reward, or removal.

Florida's current system does utilize a combination of VAM scores and discriminatory observational systems; however, whether this system is equitable in implementation is the subject of further analysis. Unlike reading and mathematics assessments, data is limited at the eighth grade level for FCAT Science 2.0. Thus, the majority of students who take the Biology EOC at the tenth grade level have not taken a state mandated science assessment since the eighth grade. Thus, the data being used by policymakers in calculating VAM for biology teachers is subject to larger errors than their reading and mathematics counterparts. Finally, this study reveals that reading proficiency explains a significant portion of the variability in biology achievement. Although the role of the teacher was not directly addressed in this study, one can reasonably conclude that biology achievement is the product of more external factors than just teacher. The VAM model, based on this finding, would unduly penalize teachers who have students with low reading proficiency levels.

On a national scale, the study’s results have some positive and negative indicators for the future. On the positive note, the reduction in the achievement gap between genders may represent a potential increase in women entering STEM fields in the future. Although this study utilized a small sample from one high school, data reported from elsewhere in the state paints a similar picture. However, the improved achievement of one group is not representative of other subsets of the population,

particularly low socioeconomic and African American students. These groups were disproportionately represented in the bottom third of the assessment scores. If this trend continues, these groups are at a greater risk of failing this assessment required for graduation, thereby creating gaps in future opportunities. The diversity of the future science workforce will therefore continue to be compromised, reducing our nation's viability and global competitiveness while simultaneously increasing both intellectual and fiscal poverty and reducing the earning potential of a subset of the nation's workforce.

Limitations of the Study

Caution should also be exercised in interpreting the findings, particularly because of the discrepancies in the measurement of reading proficiency. While the FCAT Reading assessment was used to evaluate all students, the scores are not reflective of proficiency based on a single and identical test. In fact, although the ninth-grade students were tested, their test content varied remarkably from those enrolled at the tenth-grade level, as well as those in the eleventh-grade who may have taken the FCAT Reading test on more than one occasion. Additionally, students enrolled in grades 9 and 10 completed the newly implemented FCAT 2.0 Reading test in Spring 2012 whereas eleventh-grade students completed the FCAT Reading, an older version of the assessment based on different content standards (Florida Department of Education, 2012k) that are less demanding and rigorous (Florida Department of Education, 2012e). Since these scores were obtained from different tests measuring different content standards and may have been taken during different points of the school year, the scores are not directly comparable—although all are criterion-referenced tests in reading created by the same publisher.

Another discrepancy within the study is the conservative classification of students achievement on the Biology I end-of-course assessment. For this study, students who performed in the top third in the state were considered as proficient or having passed the Biology I end-of-course assessment; however, the state mandated achievement levels were not determined until spring 2013 (Florida Department of Education, 2012j). Additionally, the scores provided by the state for the 2013 assessment administration will be markedly different from the 2012 assessment administration. For example, the 2012 administration scores were represented using raw scores and scores derived from a T-scale range from 20 to 80, whereas the scores from the 2013 administration will be reported on a scale from 325 and 475 (Florida Department of Education, 2012l).

Additionally, due to the geographic location, demographic composition, and level of academic achievement at the specified Florida Title 1 high school, the results of the study have limited generalizability beyond the sample group and, ultimately, the school population. Of particular concern is the small sample size of Hispanic and American Indian students at the North Florida school and the limited number of eleventh grade students. Additionally, the unbalanced sample based on ethnicity and grade level also reduces the generalizability of the study to a larger population. Consequently, the study's results are not considered strong enough to inform policy, although the implications for education and research are worthwhile. Nevertheless, caution should be exercised when applying the results of this study to other settings.

Finally, the lack of control over the course placement of students served as another limitation. At the Title 1 high school, students enrolled at the honors level were more similar to each other in socioeconomic standing (high socioeconomic) than

students enrolled at the regular level. Furthermore, students at the honors level were disproportionately of White heritage whereas those enrolled at the regular level were disproportionately Black. Additionally, all students enrolled in the honors biology course at the North Florida high school were taught by the same teacher, whereas students at the regular level were taught by one of two teachers. Thus the quality and type of instruction in biology serves as a confounding variable.

Recommendations for Future Research

This study points to the need for further research on the variables that impact student achievement in biology at the secondary level. The influence of covariates on student performance necessitates more research to isolate variables in a quantifiable manner. Future research should include:

- Specific reading interventions that significantly improve student achievement in biology at the secondary level. Although most research on science achievement regarding reading interventions are broad, there is a paucity of research about biology at the secondary level. Based on the results of this study, a student's level of reading proficiency is important as it has an influence on biology proficiency. The strength and positive nature of the relationship and findings from previous studies suggest that reading interventions may serve to improve biology achievement. Current studies indicate that interventions focused on incorporating reading strategies, specifically improving background knowledge and vocabulary, should result in significant improvements in biology achievement. However, this assumption is untested at the secondary level.

- An examination of the comparability of the scores derived from a computer and paper-based administration of the Biology I EOC. Although the Florida Department of Education (2006) cited previous research regarding the comparability of biology scores attained via computer and paper-based test administrations, the current computer platform being utilized has yet to be tested with empirical research. Therefore, it is unknown whether certain student populations are disproportionately affected by the use of the computer-based testing with the existing platform and whether other factors, such as connection speed or school network constraints, have a negative impact on student performance. Empirical research is also needed to investigate biology multiple choice test items, some of which contain long passages that are timed and non-

adaptive. Although study results are inconsistent, there is a possibility that reading from a computer is different from paper-based reading; it is slower, resulting in reduced accuracy and reduced reading comprehension (Noyes & Garland, 2008). Thus it is uncertain whether the differences between computer and paper-based administrations exacerbates differences in reading proficiency. Nevertheless, the specifics regarding this type of assessment are salient, especially due to the high stakes nature of the assessment.

- An investigation of the relationship between math proficiency and biology student achievement. A previous study has indicated that reading and mathematics accounted for 58% of the variation in science achievement (Maerten-Rivera, Myers, Lee, & Penfield, 2010). However, this study was not specific to biology; therefore, it is worth investigating whether mathematics proficiency contributes specifically to biology achievement.

- A study on the degree to which institutional factors impact biology achievement. In this study, 60% of the variance in biology achievement was due to student characteristics, but whether the remaining 40% is due to institutional factors (such as teacher quality, instructional strategies, and classroom environment) is the subject of further investigation.

- Research on whether females score significantly higher when instructed by female biology teachers. Research has documented the negative impact of stereotypes on academic achievement (Appel, Kronberger, & Aronson, 2011; Hoy & Hoy, 2009; Pico & Stephens, 2012; Steele & Aronson, 1995) as well as the positive impact of counter-stereotypic images on academic performance (Good, Woodzicka, & Wingfield, 2010). Therefore, female teachers may serve as counter-stereotypic role models for female students, and may potentially improve student achievement in biology. However, research specific to the academic impact of female biology teachers compared to their male counterparts is scant.

Summary

This study investigated the influence and relationship of four variables--reading proficiency, ethnicity, socioeconomic status, and gender--on the academic performance of students in biology. Overall, it determined that reading proficiency has a significant effect on biology achievement, although reading proficiency itself is significantly influenced by ethnicity and socioeconomic status. Therefore, of the four variables investigated, one of them (reading proficiency) could be used in future research to predict academic performance in biology. It is therefore important that this variable be

taken into consideration by education officials, policymakers, and stakeholders in formulating policies, teacher evaluation tools, and assessments. Further research should focus on this area so that the specific factors that account for differences in reading proficiency and its subsequent effects on biology achievement can be investigated. Consequently, the right measures can then be applied to help all students improve their academic achievement in state-mandated biology assessments.

APPENDIX A
DISTRICT CONSENT FORM

Dear School District Personnel,

My name is Janine Bertolotti and I am a graduate student in Educational Administration and Policy at the University of Florida under the supervision of Dr. Bernard Oliver. I am conducting research on the relationship between four variables- FCAT Reading, ethnicity, socioeconomic status, and gender; and their influence on the results of the Florida Biology EOC exam.

I would appreciate permission to use information from the district files on your ninth and tenth grade students who were enrolled in Biology 1 during the 2011-2012 school year. Specifically, I would like to collect data on their gender, grade level, Biology 1 EOC scores, FCAT Reading scores, and eligibility for free and reduced-price lunch. Only my supervisor and I will have access to this personal information. All data will be coded and kept confidential; no names or any identifying information will be reported.

If you have any questions, please feel free to contact me or my faculty supervisor, Dr. Bernard Oliver, at (352) 273-4358. I will be contacting your office within the next two weeks to confirm your approval.

Sincerely,

Janine C. Bertolotti

Graduate Student in Educational Leadership

APPENDIX A
IRB CORRESPONDENCE

UF | Institutional Review Board
UNIVERSITY of FLORIDA

PO Box 112250
Gainesville, FL 32611-2250
352-392-0433 (Phone)
352-392-9234 (Fax)
irb2@ufl.edu

May 10, 2013

TO: Janine Bertolotti

FROM: Ira S. Fischler, PhD; Chair
University of Florida
Institutional Review Board 02

SUBJECT: **Exemption of Protocol #2013-U-0560**
A Comparative Analysis of Student Performance: An Investigation of the Student Characteristics that Influence Florida Biology End-of-Course Scores

SPONSOR: None

Your protocol submission has been reviewed by the Board. The Board has determined that your protocol is exempt based on the category listed below:

> *45 CFR 46.101(b(4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.*

Should the nature of your study change or if you need to revise this protocol in any manner, please contact this office before implementing the changes.

IF:dl

APPENDIX B
DISTRICT APPROVAL LETTER

May 13, 2013

Janine Bertolotti

University of Florida

Gainesville, Fl 32605

Dear Ms. Bertolotti,

This letter is to inform you that our High School has reviewed and supports your research study titled, "*An investigation of the student characteristics that influence Florida Biology end-of-course scores at a Title 1 school*." It is our understanding that the data collection will begin on May 14, 2013. We are very interested in your efforts that may help to improve our understanding of Biology achievement on the end-of-course exam.

If you have any questions or need further assistance, please do not hesitate to contact me.

Sincerely,

Principal

APPENDIX B
DATA DICTIONARY

Variable	Description	Coding
ID	Student id number	#
Gender	Gender of student	Male = 1 Female = 2
Free/reduced lunch (FRLunch)	What is the eligibility of the student regarding free/reduced lunch?	Free = 1 Reduced = 2 Ineligible = 3
Free/reduced lunch (Freereducedlunch)	Does the student qualify for free/reduced lunch?	Yes= 1 No = 2
Ethnicity	What is the ethnic background of the student?	White = 1 Black = 2 Hispanic = 3 American Indian or Alaskan native = 4
Pass or fail FCAT Reading (PFReading)	Did the student receive a grade of 3 or higher on the reading FCAT?	Yes = 1 No = 2
FCAT Reading level (RLevel)	What is the student's FCAT reading level?	1 = 1 2 = 2 3 = 3 4 = 4 5 = 5
FCAT Reading Developmental Scale Score (RDevScore)	What is the student's FCAT Reading developmental scale score?	#
Grade level (GradeLevel)	What is the student's grade level?	9 = 9 10 = 10 11 = 11
Biology EOC Level (BioLevel)	Which third did the student score on the Biology EOC?	1 = 1 2 = 2 3 = 3
Biology T- score (BioTscore)	What T scale-score did the student earn on the Biology EOC exam? (20 - 80)	#

Pass or fail BIO EOC (PFBioEoc)	Did the student score in the top third on the Bio EOC?	Yes = 1 No = 2
Regression gender (Regressgender)	How is gender coded in the regression?	Male = 0 Female = 1
Regression ineligible (RegressNOTeligible)	How are students ineligible for free and reduced-price lunch coded in the regression?	Eligible = 0 Ineligible = 1
Regression reduced lunch (Regressreducedeligible)	How are students who are eligible for reduced-price lunch coded in the regression?	Free and not eligible = 0 Reduced =1
Regression free lunch (Regressfreelunch)	How are students who are eligible for free lunch coded in the regression?	Reduced and Not eligible = 0 Free = 1
Regression White students (RegressWhite)	How are White students coded in the regression?	Non-white = 0 White = 1
Regression Black students (RegressBlack)	How are Black students coded in the regression?	Non-Black = 0 Black = 1
Regression Hispanic students (RegressHispanic)	How are Hispanic students coded in the regression?	Non-Hispanic = 0 Hispanic = 1
Regression American Indian or Alaskan native students (RegressAmInAlNa)	How are American Indian or Alaskan native students coded in the regression?	Non American Indian/Alaskan native = 0 American Indian/Alaskan native = 1
Regression minority students (Regressminority)	How are minorities coded in the regression?	Minority = 0 Non-minority = 1
Regression 10th grade students (Regress10)	How are 10th grade students coded in the regression?	9th grade = 0 10th grade = 1
Regression 11th grade students (Regress11)	How are 11th grade students coded in the regression?	9th and 10th= 0 11th grade = 1

LIST OF REFERENCES

ACT (2010). The condition of college & career readiness 2010. Retrieved from http://www.act.org/research/policymakers/cccr10/pdf/ConditionofCollegeandCareerReadiness2010.pdf

Allen, J., Gregory, A., Mikami, A., Lun, J. Hamre, B., & Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system-secondary. *School Psychology Review, 42*(1), 76-97.

American Association for the Advancement of Science (AAAS) Project 2061 (1990). *Science for all Americans.* New York, NY: Oxford University Press.

American Association for the Advancement of Science (1989). *Science for all Americans*. New York, NY: Oxford University Press.

American Educational Research Association (2006, April 9). 43.010. AERA Distinguished lecture: Presidential Session. Retrieved from http://www.softconference.com/MEDIA/WMP/260407/260407-43.010.asp [video]

American Educational Research Association (2000). Position statement on high-stakes testing. Retrieved from http://www.aera.net/AboutAERA/AERARulesPolicies/AERAPolicyStatements/PositionStatementonHighStakesTesting/tabid/11083/Default.aspx

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing.* Washington, DC: American Educational Research Association.

Anderson, K.J.B. (2011). Science education and test-based accountability: Reviewing their relationship and exploring implications for future policy. *Science Education, 96*(1), 104-129.

Appel, M., Kronberger, N., & Aronson, J. (2011). Stereotype threat impairs ability building: Effects on test preparation among women in science and technology. *European Journal of Social Psychology, 41*(7), 904-913.

Aud, S., Fox, M., & KewalRamani, A. (2010). *Status and Trends in the Education of Racial and Ethnic Groups* (NCES 2010-015). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Baharudin, R. & Luster, T. (1998). Factors related to the quality of the home environment and children's achievement. *Journal of Family Issues, 19*(4), 375-403.

Baker, E.L. (2001). Testing and assessment: A progress report. *Educational Assessment, 7*(1), 1-12.

Bankston III, C.L. & Caldas, S.J. (1998). Family structure, schoolmates, and racial inequalities in school achievement. *Journal of Marriage & the Family, 60,* 715-724.

Barrera, A., Braley, R. T., & Slate, J. R. (2010). Beginning teacher success: An investigation into the feedback from mentors of formal mentoring programs. *Mentoring & Tutoring: Partnership in Learning, 18*(1), 61-74.

Beck, I.L., McCaslin, E.S., & McKeown, M.G. (1980). *The rational and design of a program to teach vocabulary to fourth-grade students* (ERDC Publication 1980/25). Pittsburgh: University of Pittsburgh, Learning Research Development Center.

Best, R.M., Rowe, M., Ozuru, Y., & McNamara, D.S. (2005). Deep-level comprehension of science texts: The role of the reader and the text. *Topics in Language Disorders, 25*(1), 65-83.

Bhanot, R.T. & Jovanovic, J. (2009). The links between parent behaviors and boys' and girls' science achievement beliefs. *Applied Developmental Science, 13*(1), 42-59.

Biancarosa, G. & Snow, C.E. (2004). Reading Next- A Vision for Action and Research in Middle and High School Literacy: A Report from Carnegie Corporation of New York. Washington, DC: Alliance for Excellent Education.

Blank, R., Langesen, D., & Petermann, A. (2007). State indicators of science and mathematics education. Retrieved from http://programs.ccsso.org/content/pdfs/SM%2007%20report%20part%201.pdf

Blazer, C. (2012). National trends in end-of-course assessment programs. Information capsule. Volume 1202. *Miami Dade County.* Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED536522

Bleeker, M. M. & Jacobs, J. E. (2004). Achievement in math and science: Do mothers' beliefs matter 12 years later? *Journal of Educational Psychology, 96*(1), 97-109.

Borg, M. O., Plumlee, J.P., & Stranahan, H.A. (2007). Plenty of children left behind: High-stakes testing and graduation rates in Duval county, Florida. *Educational Policy, 21*(5), 695-716.

Bray, G. B., Pascarella, E. T., & Pierson, C. T. (2004). Postsecondary education and some dimensions of literacy development: An exploration of longitudinal evidence. *Reading Research Quarterly, 39*(3), 306–330.

Butchart, R. E. (2010). Black hope, white power: Emancipation, reconstruction and the legacy of unequal schooling in the US South, 1861- 1880. *Paedagogica Historica, 46*(1/2), 33-50.

Bybee, R.W. (2010). What is STEM Education? *Science*, *329(*5994*),* 996-999.

Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychological Science, 16*(7), 572- 578.

Caldas, S.J. & Bankston III, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research, 90(*5*),* 269-277.

Carnoy, M. & Loeb, S. (2002). Does external accountability affect student outcomes? A crossstate analysis. *Educational Evaluation and Policy Analysis*, *24*(4), 305-332.

Center on Educational Policy (2008, August). State high school exit exams: A move toward end-of-course exams. Retrieved from www.cep-dc.org%2Fcfcontent_file.cfm%3FAttachment%3DCEP_HSEE08Report_081308.pdf&ei=2WRsUryQNaTC2wWL94GACw&usg=AFQjCNHadAcjYxNKwPmIrle98B3GkVnxBA&sig2=GX9XJQc0lCmGLaUGLICczQ

Chambers, J.G., Levin, J.D., & Parrish, T.B. (2006). Examining the relationship between educational outcomes and gaps in funding: An extension of the New York adequacy study. *Peabody Journal of Education, 81*(2), 1-32.

Chomitz, V.R., Slining, M.M., McGowan, R.J., Mitchell, S.E., Dawson, G.F., & Hacker, K.A. (2009). Is there a relationship between physical fitness and academic achievement? Positive results from public school children in the Northeastern United States. *Journal of School Health, 79*(1), 30-37.

Christensen, C. M., Horn, M. B., & Johnson, C. W. (2011) *Disrupting Class: How Disruptive Innovation Will Change the Way the World Learns*. New York: McGraw Hill.

Coleman, J.S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfield, F., & York, R. (1966). Equality of educational opportunity. Washington, DC: U.S. Government Printing Office.

College Board (2011). 2011 College-Bound Seniors: Total group profile report. Retrieved from http://professionals.collegeboard.com/profdownload/cbs2011_total_group_report.pdf

Common Core Standards Initiative (2012a). Frequently asked questions. Retrieved from http://www.corestandards.org/resources/frequently-asked-questions

Common Core Standards Initiative (2012b). Key points in English Language Arts. Retrieved from http://www.corestandards.org/resources/key-points-in-english-language-arts

Common Core Standards Initiative (2012c). Mission Statement. Retrieved from http://www.corestandards.org.

Condron, D.J. & Roscigno, V.J. (2003). Disparities within: Unequal spending and achievement in an urban school district. *Sociology of Education, 76*(1), 18-36.

Council of Chief State School Officers (2011). State end-of-course testing programs: A policy brief. Retrieved from http://www.ccsso.org/Documents/2011/State_End-of-Course_Testing_Programs_2011.pdf

Cromley, J.G. (2009). Reading achievement and science proficiency: International comparisons from the programme on international student assessment. *Reading Psychology, 30*, 89 - 118.

Cromley, J.G., Snyder-Hogan, L.E., & Luciw-Dubas, U.A. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 102*(3), 687-700.

Darling-Hammond, L. (2006). No Child Left Behind and High School Reform. *Harvard Educational Review, 76*(4), 642–663.

Davis, C.L. & Cooper, S. (2011). Fitness, fatness, cognition, behavior, and academic achievement among overweight children: Do cross-sectional associations correspond to exercise trial outcomes? *Preventative Medicine, 52*(1), S65-S69

Demps, D. L. & Onwuegbuzie, A. J. (2001). The relationship between eighthgrade reading scores and achievement on the Georgia High School Graduation Test. *Research in the Schools, 8*(2), 1–9.

Dempster, E. R. & Reddy, V. (2007). Item readability and science achievement in TIMSS 2003 in South Africa. *Science Education, 91*(6), 906–925.

Desimone, L.M. & Long, D. (2010). Teacher effects and the achievement gap: Do teacher and teaching quality influence the achievement gap between Black and White and high- and low-SES students in the early grades? *Teachers College Record, 112*(12), 3024-3073.

Dewitt, J., Archer, L., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2011). High aspirations but low progression: The science aspirations-careers paradox amongst minority ethnic students. *International Journal of Mathematics and Science Education, 9*, 243-271.

Doppen, F. S. (2007). Pre-Service Social Studies Teachers' Perceptions of High-Standards, High-Stakes. *International Journal of Social Education, 21*(2), 18–45.

Easton-Brooks, D. & Davis, A.(2007, Fall). Wealth, traditional socioeconomic indicators, and the achievement debt. *Journal of Negro Education, 76*(4), 530-541.

Ediger, M. (2005, March). Assessing reading in the science curriculum. *College Student Journal, 39*(1), 26-30.

Education Information and Accountability Services (2009, September). Data Report: Series 2010-07D- Free/reduced-price lunch eligibility. Retrieved from http://www.fldoe.org/eias/eiaspubs/pdf/frplunch.pdf

Escudier, M.P., Newton, T.J., Cox, M.J., Reynolds, P.A., & Odell, E.W. (2011). University students' attainment and perceptions of computer delivered assessment; a comparison between computer-based and traditional tests in a 'high-stakes' examination. *Journal of Computer Assisted Learning, 27*, 440-447.

Everson, H.T. & Millsap, R.E. (2004). Beyond individual differences: Exploring school effects on SAT scores. *Educational Psychologist, 39*(3), 157-172.

Fang, Z. (2006). The Language Demands of Science Reading in Middle School. *International Journal of Science Education, 28(*5), 491-520.

Fang, Z. & Wei, Y. (2010). Improving middle school students' science literacy through reading infusion. *Journal of Educational Research, 103*, 262-273.

Feathers, K.M. (2004). *Infotext: Reading and Learning.* Toronto: Pippin Publishing.

Fensham, P.J. (2009). The link between policy and practice in science education: The role of research. *Science Education, 93*(6), 1076-1095.

Fisher, D., Ross, D., & Grant, M. (2010, Jan.). Building background knowledge. *Science Teacher, 77*(1), 23-26.

Fleischman, H.L., Hopstock, P.J., Pelczar, M.P., & Shelley, B.E. (2010). *Highlights From PISA 2009: Performance of U.S. 15-Year-Old Students in Reading, Mathematics, and Science Literacy in an International Context* (NCES 2011-004). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Florida Department of Education (2010). Next Generation Sunshine State Standards. Retrieved from http://www.fldoe.org/bii/curriculum/sss/

Florida Department of Education (2012a). Biology I End-Of-Course Assessment Next Generation Sunshine State Standards: State Report of Districts. Retrieved from http://fcat.fldoe.org/mediapacket/2012/xls/Biology1EOCStateReportDistrictsv3.xls

Florida Department of Education (2012b). Biology I End-Of-Course Assessment Statewide Demographic Results: Spring 2012- Combined forms. Retrieved from http://fcat.fldoe.org/mediapacket/2012/xls/Spr2012_Bio1EOC_State_Demographic_Report_Combined_Forms.xls

Florida Department of Education (2012c). 2012 Florida Comprehensive Assessment Test 2.0 (FCAT 2.0)- Next Generation Sunshine State Standards- State Report of District Results- Grade 05- Science. Retrieved from http://fcat.fldoe.org/mediapacket/2012/xls/2012_GR05_Sdisv2.xls

Florida Department of Education (2012d). 2012 Florida Comprehensive Assessment Test 2.0 (FCAT 2.0)- Next Generation Sunshine State Standards- State Report of District Results- Grade 08- Science. Retrieved from http://fcat.fldoe.org/mediapacket/2012/xls/2012_GR08_Sdisv2.xls

Florida Department of Education (2012e). 2012 Florida Comprehensive Assessment Test 2.0 (FCAT 2.0)- Next Generation Sunshine State Standards- State Report of District Results- Grade 10- Reading. Retrieved from http://fcat.fldoe.org/mediapacket/2012/xls/GR10_Rdis_2012v4.xls

Florida Department of Education (2012f, September). Graduation requirements for Florida's statewide assessments. Retrieved from http://fcat.fldoe.org/pdf/fcatpass.pdf

Florida Department of Education (2012g, May 24). Spring 2012 Biology 1 End-of-Course Assessment: State Summary. Retrieved from http://fcat.fldoe.org/mediapacket/2012/pdf/Biology1StateSummary.pdf

Florida Department of Education (2012h). All districts: Non-high schools. Retrieved from http://schoolgrades.fldoe.org/xls/1112/All-districts-11-12.xls

Florida Department of Education (2012i). 2012 Biology 1 End-of-Course (EOC) Assessment Fact Sheet. Retrieved from http://fcat.fldoe.org/eoc/pdf/12beocfs.pdf

Florida Department of Education- Office of Assessment (2012j). FCAT 2.0 and Florida EOC Assessments Achievement Levels. Retrieved from http://fcat.fldoe.org/fcat2/pdf/achlevel.pdf

Florida Department of Education (2012k). Fall 2012: Reading & Mathematics Retake- Test administration manual. Retrieved from http://www.pearsonaccess.com/media/fl/docs/FL00005262_F12RMRt_TAM_WT_FINAL_APPROVED.pdf

Florida Department of Education (2012l). Statement from Commissioner of Education Pam Stewart on Achievement Level Recommendations for Geometry and Biology End-of-Course Assessments and FCAT 2.0 Science. Retrieved from http://www.fldoe.org/news/2012/2012_11_19-2.asp

Florida Department of Education (2012m). Standard setting recommendations for FCAT 2.0 Science (grades 5 and 8), Biology 1 and Geometry EOC assessments. Retrieved from http://fcat.fldoe.org/pdf/StandardsImpact.pdf

Florida Department of Education (2012n).Detailed information on high schools. Retrieved from http://schoolgrades.fldoe.org/xls/1112/HighSchoolGradesAllDistricts.xls

Florida Department of Education (2012o).School, district, and state public accountability report. Retrieved from http://doeweb-prd.doe.state.fl.us/eds/nclbspar/year1112/nclb1112.cfm?dist_schl=13_7361

Florida Department of Education (2012p). School, district, and state public accountability report. Retrieved from http://doeweb-prd.doe.state.fl.us/eds/nclbspar/year1112/nclb1112.cfm?dist_schl=40_11

Florida Department of Education (2012q). Florida's value-added model: Overview of the model used to measure student learning growth on FCAT Reading and Mathematics. Retrieved from http://www.feraonline.org/presentations/FloridaValueAddedModelJuanCopaFLDOE.pdf

Florida Department of Education (2011a). Florida Comprehensive Assessment Test (FCAT) 2011- Sunshine State Standards- State Report of School Results- Grade 11- Science. Retrieved from http://fcat.fldoe.org/xls/2011/F11_GR11_Ssch.xls

Florida Department of Education (2011b). Science scores: Statewide Comparison for 2001 to 2011- FCAT Science-Sunshine State Standards Test. Retrieved from http://fcat.fldoe.org/mediapacket/2011/pdf/2011ScienceComparison.pdf

Florida Department of Education (2011c). Biology I End-of-course assessment test item specifications. Retrieved from http://fcat.fldoe.org/eoc/pdf/BiologyFL11Sp.pdf

Florida Department of Education (2011d). Florida differentiated accountability program: 2011-2012 School Improvement Plan. Retrieved from http://www.flbsi.org/1112_SIP/Public/print.aspx?uid=400011

Florida Department of Education (2010). Florida Comprehensive Assessment Test (FCAT) 2010- Sunshine State Standards- State Report of School Results- Grade 11- Science. Retrieved from http://fcat.fldoe.org/xls/2010/F10_GR11_Ssch.xls

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

Florida Department of Education (2005a). Florida End-of-Course (EOC) Assessments. Retrieved from http://fcat.fldoe.org/eoc/

Florida Department of Education (2005b). Information about Florida's State of Board of Education Strategic Plan 2012- 2018. Retrieved from http://www.fldoe.org/Strategic_Plan/

Florida Department of State (2013). 6A-6.054 K-12 Student Reading Intervention Requirements. Retrieved from https://www.flrules.org/gateway/notice_Files.asp?ID=12969693

Florida Department of State (2011). 6A-5.090 Content Area Reading Professional Development and Next Generation Content Area Professional Development. Retrieved from https://www.flrules.org/gateway/notice_Files.asp?ID=9823789

Florida Legislature (2012). The 2012 Florida Statutes: 1008.22 Student assessment program for public schools. Retrieved from http://www.leg.state.fl.us/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=1000-1099/1008/Sections/1008.22.html

Florida Legislature (2010). 2010 Florida Statutes (including Special Session A). 1012.34 Assessment procedures and criteria. Retrieved from http://www.flsenate.gov/laws/statutes/2010/1012.34

Frein, S.T. (2011). Comparing in-class and out-of-class computer-based tests to traditional paper-and-pencil tests in introductory psychology courses. *Teaching of Psychology, 38*(4), 282-287.

Gay, L.R., Mills, G.E., & Airasian, P. (2009*). Educational Research: Competencies for Analysis and Applications*. Upper Saddle River, NJ: Pearson.

George, R. (2006). A cross-domain analysis of change in students' attitudes toward science and attitudes about the utility of science. *International Journal of Science Education, 28*(6)*,* 571-589.

Gibson, H. L. & Chase, C. (2002). Longitudinal impact of an inquiry- based science program on middle school students' attitudes toward science. *Science Education*, *86*(5), 693-705.

Golden, D. (2003, December 29). Initiative to leave No Child Behind leaves out gifted: Educators divert resources from classes for smartest to focus on basic literacy. *The Wall Street Journal,* Retrieved from http://online.wsj.com/public/resources/documents/Polk_Gifted.htm

Good, J.J., Woodzicka, J.A., & Wingfield, L.C. (2010). The effects of gender stereotypic and counter-stereotypic textbook images on science performance. *Journal of Social Psychology, 150*(2), 132-147.

Goodwin, B., Englert, K., & Cicchinelli, L. (2003). *Comprehensive accountability systems: A framework for evaluation.* Aurora, CO: Mid-continent Research for Education and Learning.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and Science Achievement of U.S. Fourth- and Eighth-Grade Students in an International Context* (NCES 2009-001 Revised). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Green III, P.C., Baker, B.D., & Oluwole, J. (2012). The legal and policy implications of value-added teacher assessment policies. *Brigham Young University Education & Law Journal, 1,* 1-29.

Greenfield, T. A. (1997). Gender- and grade-level differences in science interest and participation. *Science Education, 81*(3)*,* 259-276.

Gregory, A. & Weinstein, R.S. (2008). The discipline gap and African Americans: Defiance or cooperation in the high school classroom. *Journal of School Psychology, 46*(4), 455-475.

Grissmer, D. W., Kirby, S. N., Berends, M., & Williamson, S. (1994). *Student achievement and the changing American family: An executive summary.* Retrieved from the Rand on-line database on the World Wide Web: http://www.rand.org/publications/MR/MR535/MR535.html

Grossman, P.L. & Stodolsky, S.S. (1995). Content as context: The role of school subjects in secondary school teaching. *Educational Researcher, 24(*8*),* 5-11.

Gurley-Calvez , T. & Higginbotham, A. (2010). Childhood obesity, academic achievement, and school expenditures. *Public Finance Review, 38*(5), 619-646.

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., Scafiddi, N.T, & Tonks, S. (2004). Increasing reading comprehension and engagement through Concept-Oriented Reading Instruction. *Journal of Educational Psychology, 96*(3), 403–423.

Haladyna, T., Nolen, S., & Haas, N. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher, 20*(5), 2-7.

Hanushek, E.A. & Raymond, M.E. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management, 24*(2), 297-327.

Harmon, J.M., Hedrick, W.B., & Wood, K.D. (2005). Research on vocabulary instruction in the content areas: Implications for struggling readers. *Reading & Writing Quarterly, 21,* 261-280.

Harrison, C. & Cohen-Vogel, L. (2012). The politics of teacher reform in Florida: Analyzing causal narrative surrounding state adoption of performance-based evaluations, performance pay, and tenure elimination. *Peabody Journal of Education, 87,* 517-534.

Haught, P. A. & Walls, R. T. (2004). Relationships of reading, MCAT, and USMLE Step 1 Test results for medical students. *Reading Psychology: An International Quarterly, 25*(2), 83– 92.

Haworth, C.M.A., Dale, P.S., & Plomin, R. (2009). Sex differences and science: The etiology of science excellence. *Journal of Child Psychology and Psychiatry, 509*(9), 1113-1120.

Hidi, S. & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research*, *70*(2), 151-179.

Hill, H.C., Kapitula, L., & Umland, K. (2011). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal, 48*(3), 794-831.

Hines, E. & Holcomb-McCoy, C. (2013). Parental characteristics, ecological factors, and the academic achievement of African American males. *Journal of Counseling & Development, 91(*1*),* 68-77.

Hollar, D., Messiah, S.E., Lopez-Mitnik, G., Hollar, L., Almon, M. & Agatston, A.S. (2010, April). Effect of a two-year obesity prevention intervention on percentile changes in body mass index and academic performance in low-income elementary school children. *American Journal of Public Health, 100*(4), 646- 653.

Homan, S., Hewitt, M., & Linder, J. (1994). The development and validation of a formula for measuring single sentence test item readability. *Journal of Educational Measurement, 31*(4), 349 - 358.

Horne, J. (2007). Gender differences in computerised and conventional educational tests. *Journal of Computer Assisted Learning, 23,* 47-55.

House, J.S. & Williams, D.R. (2000). "Understanding and Reducing Socioeconomic and Racial/Ethnic Disparities in Health." Pp. 81-124 in *Promoting Health: Intervention Strategies From Social and Behavioral Research*, Eds. B. D. Smedley and S. L. Syme . Washington, D.C.: National Academy of Sciences Press.

Hoy, A. W. & Hoy, W. K. (2009). *Instructional leadership: A research-based guide to learning in schools* (3rd ed.). Boston: Allyn and Bacon.

Johnson, C.C. (2013). Educational turbulence: The influence of macro and micro-policy on science education reform. *Journal of Science Teacher Education, 24*(4), 693-715.

Kahlenberg, R.D. (2001). *All together now: Creating middle-class schools through public school choice.* Washington, DC: Brookings Institution.

Kendeou, P. & Van Den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition, 35*(7), 1567-1577.

Kingston, N.M. (2009). Comparability of computer-and paper-administered multiple-choice tests for K-12 populations: A synthesis. *Applied Measurement in Education, 22*, 22-37.

Konstantopoulos, S. & Chung, V. (2011). Teacher effects on minority and disadvantaged students' grade 4 achievement. *The Journal of Educational Research, 104*, 73-86.

Kress, S., Zechmann, S., & Schmitten, J. (2011). When performance matters: The past, present, and future of consequential accountability in public education. *Harvard Journal On Legislation*, *48*(1), 185-234.

Langer, J.A. (2001). Approaches toward meaning in low- and high-rated readers. Retrieved June 11, 2007, from http://cela.albany.edu/reports/langer/langerapproaches.pdf

Lee, V. E., Bryk, A. S., & Smith, J. B. (1993). The organization of effective high schools. In L. Darling-Hammond (Ed.), *Review of Research in Education, 19,* pp. 171-267.Washington, DC: American Educational Research Association.

Lesley, M. (2004). Looking for critical literacy with postbaccalaureate content area literacy students. *International Reading Association, 48(*4*),* 320- 334.

Liu, X. & Lu, K. (2008, September). Student performance and family socioeconomic status: Results from a survey of compulsory education in western China. *Chinese education and society, 41(*5), 70-83.

Long, J.S. & Fox, M.F. (1995). Scientific careers: Universalism and particularism. *Annual Review of Sociology,* 21, 45-71.

Lovelace, S. & Wheeler, T. (2006). Cultural discontinuity between home and school language socialization patterns: Implications for teachers. *Education, 127*(3), 303-309.

Madaus, G. & Russell, M. (2010). Paradoxes of high-stakes testing. *Journal of Education, 190(*1/2*),* 21-30.

Maerten-Rivera, J., Myers, N., Lee, O., & Penfield, R. (2010). Student and school predictors of high-stakes assessment in science. *Science Education, 94,* 937-962.

Marchant, G. (2004). What is at stake with high stakes testing? A discussion of issues and research. *Ohio Journal of Science, 104*(2), 2-7.

Marzano, R. (2004). *Building Background Knowledge for Academic Achievement: Research on What Works in Schools*. Alexandria, VA: Association for Supervision & Curriculum Development.

Marzano, R. & Pickering, D. (2005). *Building academic vocabulary: Teacher's manual.* Alexandria, VA: ASCD

Mattison, E. & Aber, M.S. (2007). Closing the achievement gap: The association of racial climate with achievement and behavioral outcomes. *American Journal of Community Psychology, 40,* 1-12.

McIntosh, S. (2012). State high school exit exams: A policy in transition. *Center on Education Policy.* Retrieved from http://www.cep-dc.org/cfcontent_file.cfm?Attachment=McIntosh%5FReport%5FHSEE2012%5F9%2E19%2E12%2Epdf

Mead, A. D. & Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive ability tests: A meta-analysis. Psychological Bulletin, *114*(3), 449-58.

Medina, M. & Mishra, S. (1994). Relationships among Spanish reading achievement and selected content areas for fluent- and limited-Spanish-proficient Mexican Americans. *Bilingual Review/Revista Bilingue, 19*(2), 134–141.

Meltzer, J.(2001). *Supporting Adolescent literacy across the content areas: Perspectives on policy and practice.* Retrieved from http://www.alliance.brown.edu/pubs/perspectives/adlitcontent.pdf

Meltzer, J., Cook Smith, N., & Clark, H. (2002). Adolescent literacy resources: Linking research and practice. Retrieved from http://www.alliance.brown.edu/pubs/adlit/alr_lrp.pdf

Moore, G.W. & Slate, J.R. (2008). Who's taking the advanced placement courses and how are they doing: A statewide two-year study. *The High School Journal, 92*(1), 56-67.

Mostafa, T. (2010). Decomposing inequalities in performance scores: The role of student background, peer effects and school characteristics. *International Review of Education, 56*(5/6), 567-589.

Murphy, P., DeArmond, M., & Guin, K. (2003). A national crisis or localized problems? Getting perspective on the scope and scale of the teacher shortage. *Education Policy Analysis Archives*, *11*(23). Retrieved from http://epaa.asu.edu/ojs/article/download/251/377

Nasr, A.R. & Soltani K., A. (2011). Attitude towards Biology and its effects on student's achievement. *International Journal of Biology, 3*(4), 100-104.

National Center for Children in Poverty (2007). Parents' low education leads to low income, despite full-time employment. Retrieved from http://www.nccp.org/publications/pdf/text_786.pdf

National Center for Education Statistics (2012). *The Nation's Report Card: Science 2011* (NCES 2012-465). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

National Center for Education Statistics (2011). *The Nation's Report Card: Reading 2011* (NCES 2012-457). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

National Commission on Excellence in Education (1983). *A Nation At Risk: The Imperative for Educational Reform*. Washington, DC: Government Printing Office.

National Commission on Mathematics and Science Teaching for the 21st Century (2000). *Before It's Too Late: A Report to the Nation from The National Commission on Mathematics and Science Teaching for the 21st Century.* Washington, DC: U.S. Department of Education.

National Governors Association Center for Best Practices, Council of Chief State School Officers (2010). Common Core Standards. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers.

National Research Council (2005). *America's lab report: Investigations in high school science*. Washington, DC: National Academies Press.

National Science Board (2012). *Science and Engineering Indicators 2012*. Retrieved from http://www.nsf.gov/statistics/seind12/c1/c1h.htm

National Science Board (2004, May). *Science and Engineering Indicators 2004*. Retrieved from http://www.nsf.gov/statistics/seind04/toc.htm

National Science Foundation (2012a). Science and engineering indicators 2012. Retrieved from http://www.nsf.gov/statistics/seind12/

National Science Foundation (2012b). WebCASPAR- Integrated Science and Engineering Resources Data System- Table builder: IPEDS Completions Survey by Race. Retrieved from https://caspar.nsf.gov./EmailedTable?table=1016121345202757143

National Science Foundation (2012c). WebCASPAR- Integrated Science and Engineering Resources Data System- Table builder: NSF-NIH Survey of Graduate Students & Postdoctorates in Science and Engineering. Retrieved from https://caspar.nsf.gov./EmailedTable?table=1016121334032321826

National Science Foundation, Division of Science Resources Statistics (2011). *Women, minorities, and persons with disabilities in science and engineering: 2011*.

Special Report NSF 11-309. Arlington, VA. Retrieved from http://www.nsf.gov/statistics/wmpd/pdf/wmpd2011.pdf

Neufeld, P. (2005). Comprehension instruction in content area classes. *The Reading Teacher, 59*(4), 302-312.

No Child Left Behind Act of 2001, 20 USC § 6301 (2002). Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

Nolen, S. B. (2003). Learning environment, motivation, and achievement in high school science. *Journal of Research in Science Teaching, 40*(4), 347–368.

Norris, S. P. & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education, 87*(2), 224–240.

Nosek, B.A., Smyth, F.L., Sriram, N., Lindner, N.M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Neto, F., Olli, E., Park, J., Schnabel, K., Shiomura, K., Tudor Tulbure, B., Wiers, R.W., Somogyi, M., Akrami, N., Ekehammar, B., Vianello, M., Banaji, M.R., & Greenwald, A.G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *PNAS, 106*(26), 10593-10597.

Noyes, J.M. & Garland, K.J. (2008). Computer-vs. paper-based tasks: Are they equivalent? *Ergonomics, 51*(9), 1352- 1375.

Ogbu, J. U. & Simons, H. D. (1998). Voluntary and Involuntary minorities: A cultural ecological theory of school performance with some implications for education. *Anthropology and Education Quarterly, 29*(2), 155-188.

Okpala, C.O., Smith, F., Jones, E., & Ellis, R. (2000). A clear link between school and teacher characteristics, student demographics, and student achievement. *Education, 120*(3), 487-494.

O'Reilly, T. & McNamara, D. S. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. *American Educational Research Journal, 44*(1), 161–196.

Organisation for Economic Co-operation and Development. (2007). PISA 2006: Science competencies for tomorrow's world. Paris: Author.

Organisation for Economic Co-operation and Development (OECD). (2002a). Measuring student knowledge and skills: The PISA 2000 assessment of reading, mathematical and scientific literacy. Paris: Author.

Organisation for Economic Co-operation and Development (OECD). (2002b). PISA 2000 technical report. Paris: Author.

Ou, D. (2009). To leave or not to leave? A regression discontinuity analysis of the impact of failing the high school exit exam. Economics of Education Review, *29*, 171-186.

Ozuru, Y., Dempsey, K., & McNamara, D.S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction*, *19*, 228-242.

Papay, J.P., Murnane, R.J., & Willet, J.B. (2010). The consequences of high school exit examinations for low-performing urban students: Evidence from Massachusetts. *Educational Evaluation and Policy Analysis*, *32*(1), 5-23.

Pearson Education (2013). Script for administering the Biology I EOC Practice Test (ePAT). Retrieved from http://www.pearsonaccess.com/cs/Satellite?blobcol=urlblob&blobheader=application%2Fforcedownload&blobheadername1=ContentDisposition&blobheadername2=ContentTransferEncoding&blobheadervalue1=attachment%3B+filename%3D%222012+Biology+1+EOC+TestNav+ePAT+Script_WT_FINAL.pdf%22&blobheadervalue2=binary&blobkey=id&blobtable=BinaryData&blobwhere=1255049960855&ssbinary=true

Penfield, R.D. & Lee, O. (2010). Test-based accountability: Potential benefits and pitfalls of science assessment with student diversity. *Journal of Research in Science Teaching, 47*(1), 6-24.

Perie, M., Grigg, W., & Donahue, P. (2005). The Nation's Report Card: Reading 2005. (NCES 2006–451). U.S. Department of Education, National Center for Education Statistics. Washington, D.C.: U.S. Government Printing Office. Retrieved from http://nces.ed.gov/nationsreportcard/pdf/main2005/2006451.pdf

Perry, L. & McConney, A. (2010). School socio-economic composition and student outcomes in Australia: Implications for educational policy. *Australian Journal Of Education (ACER Press)*, *54*(1), 72-85.

Pico, K. & Stephens, J.M. (2012). Culture, context and stereotype threat: A comparative analysis of young Ugandan women in coed and single-sex schools. *The Journal of Educational Research, 105*, 52-63.

Pinder, P.J. (2012). Afro-Caribbean and African American students, family factors, and the influence on science performance in the United States: The untold story. *Education, 132*(4), 725-738.

Pomplum, M., Ritchie, T., & Custer, M. (2006). Factors in paper-and-pencil and computer reading score differences at the primary grades. *Educational Assessment, 11*(2), 127-143.

Pong, S.L., Dronkers, J., & Hampden-Thompson, G. (2003). Family policies and children's school achievement in single- versus two-parent families. *Journal of Marriage and Family, 65,* 681-699.

Prokop, P., Prokop, M., & Tunnicliffe, S. D. (2007, Winter). Is biology boring? Student attitudes toward biology. *Journal of Biological Education, 42*(1), 36-39.

Puma, M. J., Karweit, N., Price, C., Ricciuti, A., Thompson, W., & Vaden-Kiernan, M. . (1997, April). *Prospects: Final report on student outcomes* (Washington, DC: U.S. Department of Education). Retrieved from http://www.eric.ed.gov/PDFS/ED413411.pdf

Qiu, X. & Wu, S. (2011). Global and local regression analysis of factors of American College Test (ACT) score for public high schools in the state of Missouri. *Annals of the Association of American Geographers, 101*(1), 63-83.

Rathbun, A. & West, J. (2004) *From Kindergarten Through Third Grade: Children's Beginning School Experiences* (NCES 2004–007). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Reeves, D.B. (2000). The 90/90/90 Schools: A case study. In *Accountability in action: A blueprint for learning organizations, 2nd ed*. (pp. 185-208). Denver, CO: Advanced Learning Press.

Riegle-Crumb, C., Farkas, G., & Muller, C. (2006). The role of gender and friendship in advanced course taking. *Sociology of Education, 79*(3), 206-228.

Riegle-Crumb, C., Moore, C., & Ramos-Wada, A. (2010). Who wants to have a career in science or math? Exploring adolescents' future aspirations by gender and race/ethnicity. *Science Education, 95*(3), 458- 476.

Roe, B. D., Stoodt, B. D., & Burns, P. C. (1991). *Secondary school reading instruction: The content areas.* Boston: Houghton Mifflin.

Romance, N. R. & Vitale, M. R. (2008, March). *Science IDEAS: A knowledge-based model for accelerating reading/literacy through in-depth science learning.* Paper presented at the annual meeting of the American Educational Research Association, New York.

Rumberger, R.W. & Palardy, G.J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record, 107*(9),1999-2045.

Russell, M. (1999). Testing on computers: a follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives, 7*(20). Retrieved from http://epaa.asu.edu/epaa/v7n20

Russell, M. & Haney, W. (2000). Bridging the gap between testing and technology in schools. *Education Policy Analysis Archives, 8*(19). Retrieved from http://epaa.asu.edu/epaa/v8n19.html

Russell, M. & Haney, W. (1997). Testing writing on computers: an experiment comparing student performance on tests conducted via computer and via paper . *Educational Policy Analysis Archives, 5*(3). Retrieved from http://epaa.asu.edu/epaa/v5n3.html

Ryabov, I. & Van Hook, J. (2007). School segregation and academic achievement among Hispanic children. *Social Science Research*, *36*(2), 767-788.

Ryan, J.E. (1999). The influence of race in school finance reform. *Michigan Law Review, 98,* 432- 481.

Sanders, W. L. & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future students academic achievement*. Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center. Retrieved from http://www.westendweb.com/robert/CS302/Sanders.pdf

Shymansky, J. A., Yore, L. D., & Anderson, J. O. (2004). Impact of a school district's science reform effort on the achievement and attitudes of third- and fourth-grade students. *Journal of Research in Science Teaching, 41*(8), 771–790.

Sikora, J. & Pokropek, A. (2012). Gender segregation of adolescent science career plans in 50 countries. *Science Education, 96*(2), 234-264.

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *Journal of Educational Research, 95*(6), 323-332.

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *Urban Review, 34,* 317-342.

Smith, M.C. (2000). The real-world reading practices of adults. *Journal of Literacy Research, 32*(1), 25–52.

Sonnert, G. & Fox, M.F. (2012). Women, men, and academic performance in science and engineering: The gender differences in undergraduate grade point averages. *Journal of Higher Education, 83*(1), 73 -101.

Sonnert, G., Fox, M.F., & Adkins, K. (2007). Undergraduate women in science and engineering: Effects of faculty, fields, and institutions over time. *Social Science Quarterly*, *88*, 1333-1356.

Snyder, T.D., Dillow, S.A., & Hoffman, C.M. (2008). Digest of Education Statistics 2007 (NCES 2008-022). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Stansfield, W.D. (2011). Educational curriculum standards & standardized educational tests: Comparing apples & oranges. *The American Biology Teacher, 73*(7), 389-393.

Steele, C.M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797-811.

Stewart, E.B. (2007). Individual and school structural effects on African American High School Students' Academic Achievement. *High School Journal, 91(2), 16-34.*

Strand, S. (2012). The White British-Black Caribbean achievement gap: Tests, tiers and teacher expectations. *British Educational Research Journal, 38*(1), 75-101.

Sykes, G., Schneider, B., & Plank, D.N. (2009). Handbook of education policy research. New York, NY: Routledge.

Taltont, E. L. & Simpson, R. D. (1985). Relationships between peer and individual attitudes toward science among adolescent students. *Science Education, 69,* 19-24.

Thompson, D., Wood, R. C., & Crampton, F.E. (2008). *Money & Schools,* 4th ed. Larchmont, NY: Eye on Education.

Torgesen, J.K. (2006). Adolescent Literacy: An Overview of Skill and Knowledge Requirements. Retrieved from http://www.centeroninstruction.org/files/Adolescent%20Literacy%20An%20overview%20of%20skill%20and%20knowledge%20requirements.pdf

Trefil, J. (2008). Science education for everyone: Why and what? *Liberal Education,* 6-11.

Trefil, J. (2007). *Why science?* New York: Teachers College Press.

Tulkin, S.R. & Kagan, J. (1972). Mother-child interaction in the first year of life. *Child Development, 43*(1), 31-41.

United Nations Children's Fund (UNICEF) (2012, May). Measuring child poverty: New league tables of child poverty in the world's rich countries. Retrieved from http://www.unicef-irc.org/publications/pdf/rc10_eng.pdf

United States Census Bureau (2012, May 17). Most children younger than age 1 are minorities, Census Bureau reports. Retrieved from http://www.census.gov/newsroom/releases/archives/population/cb12-90.html

United States Census Bureau, Population Division (2012). *Table 5. Cumulative estimates of the components of resident population change by race and Hispanic origin for the United States: April 1, 2000 to July 1, 2009 (NC-EST2009-05)*. Retrieved from http://www.census.gov/compendia/statab/2012/tables/12s0005.pdf

United States Commission on National Security/21st Century (2001b, May 15). *Journeys through the teacher pipeline: Recapitalizing American education- Partial cost estimates for road map for national security*. Retrieved from http://www.au.af.mil/au/awc/awcgate/nssg/addendum/Education_Addendum.pdf

United States Commission on National Security/21st Century (2001a, February 15). *Road map for national security: Imperative for change- The phase III report of the U.S. Commission on National Security/21st century.* Retrieved from http://www.au.af.mil/au/awc/awcgate/nssg/phaseIIIfr.pdf

United States Congress (2002). Public Law 107-110. No Child Left Behind Act of 2001. Washington, D.C. Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/index.html

United States Department of Agriculture (2012, August). National School Lunch Program. Retrieved from http://www.fns.usda.gov/cnd/lunch/AboutLunch/NSLPFactSheet.pdf

United States Department of Education (2004, February 10). Proven methods: The facts about...science achievement. Retrieved from http://www2.ed.gov/nclb/methods/science/science.html

United States Department of Education (1989). *Report on the National Education Goals Panel*. Washington, DC: Government Printing Office.

United States Department of Education, National Center for Education Statistics (2012). *Digest of Education Statistics, 2011* (NCES 2012-001). Retrieved from http://nces.ed.gov/fastfacts/display.asp?id=171

United States Department of Labor (2010). Highlights of Women's Earnings in 2009. Retrieved from http://www.bls.gov/cps/cpswom2009.pdf

Velayutham, S., Aldridge, J.M., & Fraser, B. (2012). Gender differences in student motivation and self-regulation in science learning: A multi-group structural equation modeling analysis. *International Journal of Science and Mathematics Education, 10*, 1347-1368.

Velde Pederson, P. (2007). What is measured is treasured: The impact of the No Child Left Behind Act and nonassessed subjects. *The Clearing House, 80(*6*),* 287- 291.

Vincent, C.G., Tobin, T.J., Hawken, L.S., & Frank, J.L. (2012). Discipline referrals and access to secondary level support in elementary and middle schools: Patterns

across African-American, Hispanic-American, and White students. *Education and Treatment of Children, 35*(3), 431-458.

Visone, J. (2010). Science and reading: What is being measured by standardized tests? *American Secondary Education, 39*(1), 95- 112.

Visone, J. (2009). The validity of standardized testing in science. *American Secondary Education, 38*(1), 46-61.

Vitale, M.R. & Romance, N.R. (2012). Using in-depth science instruction to accelerate student achievement in science and reading comprehension in grades 1-2. *International Journal of Science and Mathematics Education, 10,* 457- 472.

Walton, G.M. & Spencer, S.J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science, 20*(9), 1132-1139.

White House, Office of the Press Secretary. (2011, January 25). Remarks by the President in State of Union Address [Press Release]. Retrieved from http://www.whitehouse.gov/the-press-office/2011/01/25/remarks-president-state-union-address

White House, Office of the Press Secretary. (2010, October 18). President Obama to host White House Science Fair [Press Release]. Retrieved from http://www.whitehouse.gov/the-press-office/2010/10/18/president-obama-host-white-house-science-fair

Wiliam, D. (2010). Standardized Testing and School Accountability. *Educational Psychologist*, *45*(2), 107-122

Wingfield, A. & Guthrie, J.T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology, 89*(3), 420-432.

Winstead, L. (2011). The impact of NCLB and accountability on social studies: Teacher experiences and perceptions about teaching social studies. *The Social Studies, 102,* 221- 227.

Wood, K. D. (2002). Status of high school Biology teaching. Retrieved from http://2000survey.horizon-research.com/reports/high_biology/high_biology.pdf

Woods, T.A., Kurtz-Costes, B., & Rowley, S.J. (2005). The development of stereotypes about the rich and poor: Age, race, and family income differences in beliefs. *Journal of Youth and Adolescence, 34*(5), 437-445.

Yore, L. D., Bisanz, G. L., & Hand, B. M. (2003). Examining the literacy component of science literacy: 25 years of language arts and science research. *International Journal of Science Education, 25*(6), 689–725.

Zegiob, L.E. & Forehand, R. (1975). Maternal interactive behavior as a function of race, socioeconomic status, and sex of the child. *Child Development, 46*(2), 564-568.

Zinth, J.D. (2012). End-of-Course Exams. *Education Commission of the States.* Retrieved from http://www.ecs.org/cl

BIOGRAPHICAL SKETCH

Janine Cecelia Bertolotti was born in 1983 in Kingston, Jamaica. She attended Covenant Christian Academy before enrolling at Sts. Peter and Paul Preparatory School. Subsequently, she enrolled at Immaculate Conception High School in Jamaica before matriculating at Wesleyan College in Macon, Georgia at the tender age of 17. In May 2005, she graduated *summa cum laude* with a major in psychology and double minors in biology and neuroscience. After working for approximately one year at a technical college in Georgia, she enrolled full-time in the Master's degree program in curriculum and instruction at the University of Central Florida in Orlando, Florida. Before graduating in December 2007, she served as a graduate research assistant before beginning her first year teaching biology at a public, Miami high school.

In the summer of 2010, she began a part-time doctoral program while working full-time as a biology high school teacher. She graduated in August 2012 with an Education Specialist degree in educational leadership from the University of Florida before attaining her Doctor of Education degree in educational leadership in December 2013. As a U.S. Air Force wife, she moves frequently with her husband and daughter.