Spring 2012

The Value of NJ School Distict Demographic Data in Explainining School District NJ ASK Grade 3 Language Arts and Mathematics Scores

Peter G. Turnamian
Seton Hall University

The Value of NJ School District Demographic Data in Explaining School District NJ ASK Grade 3 Language Arts and Mathematics Scores

Peter G. Turnamian

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Education

Seton Hall University

Dissertation Committee

Christopher H. Tienken, Ed.D,, Mentor
Barbara Strobert, Ed.D.
Barbara Weller, Ed.D.
Maryrose Caulfield Sloan, Ed.D.

2012

SETON HALL UNIVERSITY
COLLEGE OF EDUCATION AND HUMAN SERVICES
OFFICE OF GRADUATE STUDIES

APPROVAL FOR SUCCESSFUL DEFENSE

Doctoral Candidate, **Peter Turnamian,** has successfully defended and made the required modifications to the text of the doctoral dissertation for the **Ed.D.** during this **Spring Semester 2012**.

DISSERTATION COMMITTEE
(please sign and date beside your name)

Mentor:
Dr. Christopher Tienken Jan. 25, 2012

Committee Member:
Dr. Barbara Strobert Jan. 25, 2012

Committee Member:
Dr. Maryrose Caufield-Sloan 1/30/12

Committee Member:
Dr. Barbara Weller 1/25/12

External Reader:

The mentor and any other committee members who wish to review revisions will sign and date this document only when revisions have been completed. Please return this form to the Office of Graduate Studies, where it will be placed in the candidate's file and submit a copy with your final dissertation to be bound as page number two.

ABSTRACT

This study examined the strength and direction of the relationship between 2009 NJ ASK 3 Language Arts and Mathematics scores and district social and demographic data (i.e., lone-parent household, level of parental education, and household income levels) found in the extant literature to influence student achievement on high-stakes standardized assessments. Analysis included stepwise multiple linear regression, simultaneous multiple linear regression, and hierarchical linear regression. This study looked at the entire population of New Jersey school districts with at least 25 students enrolled in third grade in 2009. The results of this study revealed that 60 percent of school district 2009 NJ ASK 3 Math scores could be predicted within 10 points by looking solely at three out-of-school district community variables. The results of this study also revealed 52 percent of school district 2009 NJ ASK 3 Language Arts scores could be predicted within 10 points by looking solely at the same three out-of-school district community variables applied to predict 2009 NJ ASK 3 Math scores.

ACKNOWLEDGMENTS

To Dr. Tienken, thank you for your guidance and mentorship over the past two years. I enrolled in this doctoral program at a crossroads in my career. Due in large part to your teaching, this program reconnected me with my progressive roots. Thanks to you, I discovered that a large research base exists for principles that have driven my work for the past twenty years. You brought my thinking about teaching and learning to a higher level and gave me greater confidence in my value as an educator. Your passionate articulation of a large research base inspired me anew to continue rejecting the "kool-aid." I am also grateful for the patience, generosity, and kindness you showed me throughout this process. I hope we will stay in touch.

To Dr. Caulfield, thank you for your mentorship. At the start and midway point of this program, my frustration with the politics of education was at an all-time high; your wisdom and insight inspired me to stay the course. Today, you serve as a role model I aspire to emulate.

To Dr. Strobert, thank you for your generosity and kindness. It has been one of the great privileges of my career to collaborate with you. In fact, you were the first person to connect me with the Seton Hall program. I will always be grateful to you for taking an interest in me during my transition to Watchung School. You represent the very best of our profession, and I have learned much from our talks.

To Dr. Weller, thank you for your mentorship and guidance. During my transition to Watchung School you kindly made time to meet with me and within minutes I knew we shared similar values about teaching and learning. You taught me much these

past two years and also represent a role model I aim to emulate. I am forever grateful to you for serving on this
committee because I know that many other responsibilities require your attention.

To Dr. Sloan-Caulfield, thank you for agreeing to serve on this committee. As with Dr. Weller, I understand the enormity of other responsibilities requiring your attention and remain humbled by your service on my committee. It has been an honor to work with you.

To my mother and father, I could not have completed this work without the love and support you provided my family and me. Knowing that the children were being cared for and nurtured and that Erin was not alone while I was in class or writing gave me the peace of mind to complete this degree. I am eternally grateful for your love and support.

To my wife, Erin, this dissertation would not have been possible without your heroic efforts. During this two-year period, we welcomed a third child into the world and at no point did your commitment to this work or to me ever waiver. Even as I questioned the pursuit, you never did. I remain in awe of you and all you have accomplished. Nothing gives me more joy than our time together and the single greatest challenge in this work was the time it took me away for you and our family. Your understanding and love during this time meant the world to me. In my eyes, this degree includes your name.

On a lighter note, I would also like to acknowledge The Zac Brown Band, whose tunes gave me a sense of calm and joy during extended periods of writing reminiscent of my younger years and the sounds of Jimmy Buffett.

DEDICATION

To Ryan, Riley, and Abigail. You are my joy and you were the inspiration for this degree. Words cannot express my love for you, but let me not allow this shortcoming to prevent me from expressing how much my love for you impacted this work.

Over the past two years, I was driven to complete this degree in large part to model for you the values of life-long learning. Before you entered the world, I hadn't thought about continuing my education. You inspired me!

I dedicate this dissertation to you as an example of how much you can accomplish when you believe in yourself and work hard. Over the past two years, there were many moments I did not want to continue pursuing this degree. Thanks in large part to your mother's support, I persevered. During those times, you were a great source of inspiration for me; and you were always close to my heart.

As you continue to grow and discover more of our world, I hope you will develop a love of learning and a respect for work. We may be living during increasingly cynical times, but I hope you will reject the temptations of cynicism and fear and choose to see and appreciate the good around you and in others.

May you grow to be insatiably curious, love who you are and aim to be, and may you strive to work hard and be empathic toward others. May you learn to forgive yourself when you fall short of your goals and ideals AND, most important, may you always dream big!

I love you and I am proud of you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I

During the past century two dueling paradigms, Essentialism and Progressivism, have been prevalent in the evolution of America's public education system. Today, Essentialism dominates the landscape of public education with power and political support. What led to this dominance, and how did it come about?

Absent clear policies and a vision for public education, America's earliest public schools were highly simplistic and representative of individual community needs. At the start of the nineteenth century, visions for a public school system began to take shape. During this time, the work of British educator Joseph Lancaster and his ideas for how best to educate poor children gained popularity in America. Lancaster developed an educational program using a "monitorial system" to educate children. This approach involved packaged lessons being taught to large classes with an emphasis on a mechanical drill approach to teaching (Tanner & Tanner, 2007 p. 8). Plagued by debt and complaints of harsh discipline, the Lancastrian schools fell out of favor shortly after his death in 1838. Yet, their existence foreshadowed an evolving American philosophy about schooling rooted in the values of control and efficiency. With the dawn of the twentieth century and the coming industrial revolution, demand for public schools in America grew exponentially. Between 1900 and 1920 national enrollment in high schools grew from 500,000 students to over 2,000,000, and the number of high school teachers increased from 6,000 to 14,000 (Bennett, 1972). In other words, enrollment grew by 400%, while the number of teachers grew by only 130%. As a result of this dramatic increase in demand, a debate evolved about the purpose of and access to public schools in

America. This debate continues today and has produced two increasingly distinct paradigms of the American public school system.

While Lancaster was developing and selling his mechanical reforms, Johann Pestalozzi, Johann Herbart, Friedrich Frobel, and Horace Mann were offering alternative perspectives about how and what students should be taught. Together, these educators valued the "experience" (Tanner & Tanner, 2007 p. 14) as the starting point of curriculum development. Early on, Herbart went so far as to define a process for teaching that intentionally valued students' background knowledge and experience. He developed *Five Formal Steps of Teaching and Learning.* In a time when memorization and recitation were the main methods of teaching, these simple steps for instruction were well received (Tanner & Tanner, 2007 p. 15). In Horace Mann's reports to the Massachusetts Board of Education from 1837 to 1848, he consistently attacked rote learning and described learning as an active process in which the student must be the worker. Mann was one of the first to advocate for individualized education and differentiated instruction (Tanner & Tanner, 2007 p. 20). In doing so, he laid the seeds from which Progressive thought would grow to challenge the idea of mental discipline as the guiding principle for curriculum development.

William Bagley first articulated the Essentialist paradigm. He published *Essentialist's Platform* (1938) as a critique of what he believed were the extreme tendencies of American Progressivism. During his career, Bagley consistently advocated for education to be viewed as a profession, believing that teachers deserved opportunities for professional training and study. While not disagreeing with all aspects of Progressivism, he criticized Progressivism for a lack of emphasis on the sciences. He

preferred the core of all curricula be rooted in the sciences to achieve academically rigorous curricula. Bagley viewed the sciences as superior curricula and reasoned that mastery of the sciences by students would represent superior intelligence. These ideas were at odds with the Progressive values that prioritized the individual needs of the learner and liberal arts curricula (Zilversmit, 1993). While not the dominant paradigm during his career, over time theorists and school commentators like Arthur Bestor, E.D. Hirsch and Diane Ravitch expanded on Bagley's theories. These theorists rationalized that a highly structured curricula rooted in scientific fact and order causes a general improvement and training of the mind.

Aspects of the Essentialist learning process came to be known as the theory of mental discipline and later evolved into Behaviorism (Hull, 1943; Hull, 1951; Skinner, 1953). Essentialists advocated for students to be trained in certain subject matter so that they would be properly prepared for college. Interestingly, Edward Thorndike's groundbreaking research published in *Animal Intelligence* initially supported this theory of learning (Thorndike, 1911). While Thorndike's later research would demolish the theory of mental discipline and ideas about curriculum superiority, he consistently advocated an Essentialist vision of public education throughout his career, including racial and ethnic superiority (Tomlinson, 1997). He did not view intelligence as something to be enhanced or grown in all students. Instead, he interpreted his research to support the use of assessment data to rank students and then educate the superior minds toward a more perfect citizenry (Tomlinson, 1997). Over time, mental discipline theory gained increased traction and became more prevalent in curriculum initiatives serving communities of high poverty.

During President Lyndon Johnson's "War on Poverty," Essentialists Bereiter and Englemann (1966) emphasized the importance of "intensive direct instruction" and a "general bombardment for educationally disadvantaged students (Tanner & Tanner, 2007 p. 9). While not popular during the first half of the twentieth century, Essentialist theory prevails as the dominant paradigm guiding the development of present day educational policy as evidenced by NCLB and Race to the Top policies, which increasingly emphasize the importance quantifiable results, efficiency, and accountability.

Ironically, Thorndike (1924) concluded that no one course of study "is more likely than any other study to result in a general improvement of the mind (Tanner & Tanner, 2007 p. 45)." The Progressive Education Association followed Thorndike (1924) with a landmark experiential study during the 1930s, the Eight-Year Study. The Eight-Year Study added greater credence to Thorndike (1924), as well as the teaching of John Dewey and his predecessors, by demonstrating that high school graduates from experimental high school programs were not handicapped by their high school curriculum. In fact, these graduates outperformed their peers from traditional high schools (Aiken, 1942).

Francis W. Parker, often credited as the father of the American Progressive movement, was an early advocate for universal, free education for all students. Parker rejected educational theory rationalizing the need for greater standardization and rote learning. Instead, he favored the use of group activities and curricula rich in both the arts and sciences (Parker, 1894). He viewed the purpose of a public school education to be the development of the individual's mental, physical, and moral self (Tanner & Tanner, 2007). This holistic emphasis on the individual learner and de-emphasis on using grades

to rank and order students placed Parker's ideas at odds with developing Essentialist theory.

A more famous and prolific writer and Parker protégé, John Dewey, outlined the merits of and rationalization for Progressivist theory (Dewey, 1897; Dewey, 1900; Dewey, 1902; Dewey, 1916; Dewey, 1938). Central to Progressive thought, Dewey described a paradigm in which students were viewed as creators of knowledge. In 1896 Dewey established a laboratory school at the University of Chicago where he designed student-centered curricula with planned experiences to create learning for students. "Dewey developed this psychological concept into a curriculum principle: The child's impulses are an enormously important educational resource, and opportunities should be provided to children to develop the impulses through engagement in activities" (Tanner & Tanner, 2007 p. 37). The dominance of Progressive educational theory prevailed through much of the first half of the twentieth century.

Although the Eight-Year Study and Thorndike (1924) demonstrated the effectiveness of Progressive educators from Horace Mann to John Dewey, the findings did not receive much attention due to economic and political realities. Soon after the Eight-Year study America entered World War II, giving increased credence to the values of control and efficiency, as the need to militarize the nation became a priority. With the end of World War II and the arrival of the "baby boom" generation, American culture continued to evolve and so did its thinking about education. The Cold War soon took shape and the fear of "falling behind" would become a dominant and uniquely American concern. Consequently, the Essentialist tenets of discipline, control, structure, reward

and punishment became increasingly attractive to an increasingly anxious American citizenry.

The shift in dominance from Progressive theory to Essentialist theory began midway through the twentieth century. A watershed moment exemplifying this shift was America's reaction to the Soviet Union's successful launch of Sputnik I into outer space. Anxiety about the United States "falling behind" took hold within American culture and would lead to the federal government's significantly increasing its financial and policy interest in public education. The reauthorization of the 1965 Elementary and Secondary Education Act (ESEA) became one of President Johnson's landmark "War on Poverty" policy initiatives. With this increased investment of federal money in educational programs at the state level came a greater need for the federal government to quantify and measure the impact of these funds. Consequently, during the 1970s politicians began calling for the development and use of basic skill standardized tests to determine if students were achieving basic levels of minimum competency (Amrein & Berliner, March 2002).

In 1983 a newer phenomenon in the discourse about American public education took hold. A school privatization and standardized accountability movement grew after several national reports about America's public education system were published in 1983. The most significant and influential of these reports was *A Nation at Risk* (Amrein & Berliner, March 2002); (Amrein & Berliner, March 2002). This report further heightened fears concerning America's ability to remain a superpower because of a "failing" education system. Whereas the *Cardinal Principles of Secondary Education* (National Education Association of the United States, Commission on the Reorganization of

Secondary Education, 1918) can be cited as evidence for the dominance of a Progressive era in American public education during the first half of the twentieth century, *A Nation at Risk* represented the growing dominance of Essentialist theory in education policy. Soon after publication of *A Nation at Risk,* the minimum competency tests created during the 1970's were abandoned for standardized assessments aimed at increasing the rigor of the classroom experience. Through the 1980's and 1990's a series of federal policy initiatives exemplified the arrival of standards based reform culminating in the Goals 2000: Educate America Act (Goals 2000: Educate America Act (P.L. 103-227)). These policies called for new and "higher" academic standards and more standardized testing to better quantify student learning. These standardized assessments took on even greater meaning and influence when the policies attached high-stakes rewards and consequences to standardized assessment results with the enactment of No Child Left Behind (NCLB) legislation in 2002. Among the many reforms included in NCLB legislation, standardized assessments were required to be administered to all public school students in grades 3-8 annually and the results shared publicly. NCLB further increased the dominance of the Essentialist theory of American public education policy and development as evidenced by the increased emphasis placed on the development and revision of standards at all grade levels. Previously, states developed and adopted a variety of different academic standards for their K-12 public education systems. NCLB incentivized states to use standardized assessments to evaluate how aligned school and district curricula were to state's standards. NCLB further cemented Essentialist values in education policy by further defining student achievement at the school, district and national level through the use of standardized assessment aligned to a set of standards.

The next logical step in the evolution of these Essentialist policies would be the adoption of one set of standards for all states, national standards.

On June 1, 2009, the National Governors Association (NGA) and Council of Chief State School Officers (CCSS0) unveiled the Common Core State Standards. The stated purpose of these standards was to "provide a consistent clear understanding of what students are expected to learn, so teachers and parents know what they need to do to help them" (National Governors Association, 2009). On July 24, 2009, President Barack Obama and Secretary of Education Arne Duncan unveiled the Education Recovery Act as part of the American Recovery and Reinvestment Act of 2009. The Education Recovery Act included $4.35 billion in funds for the *Race to the Top Program* (RTTP). This program created incentives for states to adopt education reform policies in the following areas: great teachers and leaders, state success factors, standards and assessment (including the adoption of the Common Core Standards), general selection criteria, improvement of the lowest achieving schools, data systems to support instruction and incentives to prioritize STEM (Science, Technology, Engineering, and Math) education. Both initiatives aim to improve student achievement to ensure that every public school student graduates from high school ready for college. In less than sixty days, the landscape of American public education policy shifted still further toward Essentialism theory.

The Race to the Top Program (RTTP) also established a Comprehensive Assessment System Competition, leading to the development of two large consortiums of states: The Common Assessment Consortium facilitated by Achieve Inc. and the Florida-led Common Assessment Consortium. These consortia of states are working to design

and develop the next generation of standardized assessments aligned with the Common Core Standards. It is not yet clear if the data generated from these new assessments will be used in a high-stakes manner. However, no evidence exists to suggest that the dominance of the Essentialism paradigm, firmly in place since the mid 1980s, has begun to shift toward more Progressive theories. Therefore, it can be assumed the data generated from a new generation of assessments will be high-stakes data.

Statement of the Problem

Maylone (2002) correlated district socioeconomic data with Michigan Educational Assessment Program (MEAP) data, and through multiple regression analysis found three district socioeconomic factors combined to reliably predict a school district's composite MEAP scores (Maylone 2002). Jones (2008) found a predictive equation for New Jersey high school performance on the High School Proficiency Assessment (HSPA) using district and school demographic data published in the annual New Jersey School Report Card. Jones (2008) also compared actual New Jersey high school HSPA scores with each high school's predicted scores. This analysis allowed Jones (2008) to identify high schools she claimed were exceeding expectations, meeting expectations, or failing to meet expectations, accounting for the influence of out-of-school and in-school variables. Maylone (2002) and Jones (2008) demonstrated through multiple regression analysis of high school standardized high-stakes assessment data and district socio-economic data how a reliable predictive formula for student achievement can be created. Both studies demonstrate the value of controlling for specific socio-economic variables when analyzing high school high-stakes standardized assessment data to more accurately determine if high schools are meeting, exceeding, or failing to meet expectations. For

example, students attending a high school in an affluent community may achieve high HSPA test scores relative to similar communities. When district socioeconomic data is factored into the analysis, the same school may be shown to actually be underperforming. Maylone's (2002) findings also raise questions about the role of high-stakes standardized assessments in determining the influence of in-school variables on student achievement.

Therefore, a problem exists when policymakers use data generated from high-stakes standardized assessment to measure the quality and success of a school district and fail to accurately control for out-of-school variables present in each district's socio-economic data. Out-of-school variables such as family wealth indicators have been proven to significantly influence student achievement as measured by standardized assessments (Coleman, Hobson, McPartland, Mood, Weinfield, & York, 1966; Jencks, Smith, Acland, Bane, Cohen, Gintis, et al., 1972; Jones, 2008; Sirin, 2005; White, 1982). Also of note, state level high-stakes standardized assessments have limitations and flaws (Joint Committee on Testing Practices, 2004; National Council on Measurement in Education, 1995; Tienken, 2008(a); Tienken, 2008; Tienken, 2011; Tienken & Rodriguez, 2010).

It cannot be assumed that high-stakes data generated from standardized assessments accurately measure the quality and success of a school district. Consequently, education policymakers may be operating under the false assumption that high scores on high-stakes standardized assessments accurately identify quality and success in school districts. Education policymakers may be rewarding or punishing school districts based on a false paradigm by using high-stakes test data to identify quality and success in school districts without consideration or control for other

significant socio-economic variables proven to impact student achievement as measured by standardized assessments.

A need exists in New Jersey for empirical, quantitative analysis to determine the influence out-of-school variables such as median home income and other socioeconomic-status variables have on NJ ASK Language Arts and Mathematics scores and the predictive strength of such variables. While the influence of district socioeconomic variables has been researched to some degree at the high school level, no research about the predictive strength of district socio-economic variables at the elementary level has been conducted.

Purpose of the Study

The purpose for this study was to identify the specific school community demographic factors that account for the greatest amount of variance in a New Jersey school district's percentage of students scoring Proficient or above on NJ ASK 3 in Languages Arts and Mathematics. The study intentionally limited its focus to-out-of school variables on district NJ ASK 3 data. This limitation was set because, if out-of-school variables are found to explain significant variance and in some cases predictive power in district test scores, as the existing literature suggests, the value of using district test scores to measure the quality of in-school variables may be in question.

The study aimed to extend aspects of Maylone (2002). Where Maylone (2002) analyzed the predictive validity of district socio-economic data correlated with high-stakes high school standardized assessment data, this study focused on New Jersey elementary school district high-stakes assessment data.

No study like this has been undertaken in New Jersey since the NCLB era. Therefore a valid predictive model of district achievement data at the elementary school level could provide policymakers and school district leaders with greater insights about how best to design and implement early intervention models. Furthermore, district socioeconomic data may be proven to be a more relevant valid predictor of future high-stakes standardized assessment data than school district variables, which could also inform future policy recommendations.

Study Design and Methodology

This study used archival NJ ASK 3 school district Language Arts and Mathematics scores from 2009 and five-year estimates from U. S. Census data to determine if a predictive equation existed between the data. The grade level of student achievement examined was Grade 3 because this is the first grade in which students are administered the 2009 NJ ASK 3 in a New Jersey school district.

Research Questions

This study examined three overarching research question:

1. How much variance in NJ ASK 3 2009 test results in Language Arts and Mathematics is explained by out-of-school socioeconomic variables?

2. How accurately can out-of-school socioeconomic and community-level variables predict a school district's percentage of students scoring Proficient or above on the 2009 NJ ASK 3 Language Arts and Mathematics sections?

3. Which community-level variables account for the greatest amount of variance in a school district's percentage of students passing the 2009 NJ ASK 3?

To gain a deeper understanding about these questions, and after a thorough review of extant literature, eight research questions were developed.

Research Questions Delineated

Research Question 1: How much variance in 2009 NJ ASK 3 test results in Language Arts can be explained by the household-income construct (Table 2) for New Jersey school districts?

Research Question 2: How much variance in 2009 NJ ASK 3 test results in Mathematics can be explained by the household-income construct (Table 2) for New Jersey school districts?

Research Question 3: How much variance in 2009 NJ ASK 3 test results in Language Arts can be explained by the lone-parent household construct (Table 3) for New Jersey school districts?

Research Question 4: How much variance in 2009 NJ ASK 3 test results in Mathematics can be explained by the lone-parent household construct (Table 3) for New Jersey school districts?

Research Question 5: How much variance in 2009 NJ ASK 3 test results in Language Arts can be explained by the percentage of parental education construct (Table 4) for New Jersey school districts?

Research Question 6: How much variance in 2009 NJ ASK 3 test results in Mathematics can be explained by the percentage of parental education construct (Table 4) for New Jersey school districts?

Research Question 7: Which combination of independent variables establishes the

greatest reliable predictive power for a school district's 2009 NJ ASK 3 Language Arts test results?

Research Question 8: Which combination of independent variables establishes the greatest reliable predictive power for a school district's 2009 NJ ASK 3 Mathematics test results?

The unit of analysis for this study was the school district. The study built upon the independent variables of Maylone (2002) and included additional independent variables based on review of relevant literature.

This study examined the following independent variables:

1. Household Income, defined as:
    - Median district household income
    - Percentage of families below poverty
    - Percentage of economically disadvantaged
    - Percentage of household annual income under $30,000
    - Percentage of household annual income above $200,000
2. Lone-Parent Household, defined as:
    - Percentage of district male households, no wife
    - Percentage of district female households, no male
3. Parental Education, defined as:
    - Percentage of population 25 years or older, no high school diploma
    - Percentage of population 25 years or older, high school graduate
    - Percentage of population 25 years or older, high school graduates and some college experience

- Percentage of population 25 years or older, bachelor's degree
- Percentage of population 25 years or older, advanced degree

The dependent variables for this study were school district NJ ASK3 language arts and mathematics proficiency data, which was defined as the percentage of the student population that achieved either a "proficient" or "advanced proficient" score.

Theoretical Framework

The line of inquiry for this study aimed to establish a research base for a new policy context to better explain which out-of-school variables are predictably impacting student achievement in Grade 3 for New Jersey school districts.

Sirin (2005) showed family socioeconomic status (SES) is the most important determinant of school financing because in the United States half of all public school funding is based on the property taxes within each school district. State and federal subsidies fail to create equitable funding across school districts and communities. Based on current school financing policies, a situation is created in which students from lower SES families are most likely to attend school districts that are financially inferior to wealthier school districts.

Maylone (2002) determined the predictive power of student achievement by combining percentage of lone-parent household, mean annual district household income, and percentage of free and reduced lunch at the high school level. The literature suggests (Haveman, R. & Wolfe, 1995; Sirin, 2005) that free and reduced lunch information can be problematic as an identifier for the family effects of SES. Sirin (2005) notes that the ES (effect size) for SES influence on student achievement increases with each grade level during the primary and middle school years and then the ES decreases in high school.

This finding provides further evidence for the design of this study to focus on third grade test scores because this is where the SES influence will be in its earliest stage. A literature review is required to determine if past research has found the variables of household income, percentage of lone-parent households and the level of parental education within a school district to explain student achievement as measured by standardized tests.

Significance of the Study

Empirical data is needed to determine the predictive validity of school district socioeconomic data on student achievement as measured by high-stakes standardized assessments. It has been proven significant that out-of-school variables impact student achievement as measured by high-stakes standardized assessments (Alspaugh, 1991; Amato & Keith, 1991; Astone & McLanahan, 1991; Blau, 1999; Coleman, Hobson, McPartland, Mood, Weinfield, & York, 1966; Dawson, 1991; Downey, 1995; Hauser & Sewell, 1986; Haveman & Wolfe, 1995; Jencks, Smith, Acland, Bane, Cohen, Gintis, et al., 1972; Payne & Biddle, 1999; Peterson & Zill, 1986; Plug & Vijverberg, 2005; Roscigno & Ainsworth-Darnell, 1999; Sirin, 2005). Recent studies have shown how multiple regression analysis of district level socio-economic data and student achievement data can be used to determine a predictive formula of district level student achievement (Jones, 2008).

In New Jersey, the present system for categorizing school districts based on socio-economic data doesn't include predictive data which could be used to better determine which districts are meeting, exceeding, or failing to meet expectations while controlling for each district's socioeconomic data. Furthermore, a large body of research

demonstrates the importance of effective early childhood education and early interventions (Fryer & Levitt, 2004; Jencks & Phillips, 1998; Nagaoka & Roderick, 2005; Ramey & Ramey, 1998). Therefore, the earlier predictive valid data can be established about school districts sooner and policymakers and school district leaders can design and implement intervention strategies. In New Jersey, the earliest standardized assessment data available for study is generated from the annual administration of NJ ASK3 to all public school students at the end of their third-grade year.

Delimitations

Data for this study was gathered primarily from two sources. The 2009 NJ ASK3 data was taken from each school district's annual School Report Card along with the percentage of economically disadvantaged families. District socio-economic data was retrieved from the U.S. Census Bureau *American FactFinder*. The data was analyzed at the district level. Analysis was not conducted of individual school aggregate data.

Analysis of district socioeconomic data was delimited to those used for the socio-economic data in Maylone (2002) and variables identified from review of the literature.

The only source of student achievement for this study was 2009 NJ ASK3, since it is the only high-stakes standardized assessment administered in all New Jersey school districts.

The study looked only at New Jersey school district data. The findings of the study cannot be generalized for school districts outside of the New Jersey region. Because the unit of analysis of the study was delimited to the district level, generalizations about the findings cannot be assigned to individual schools, teachers, or student populations beyond district Grade 3. Furthermore, generalizations about the

findings at the district level cannot be assigned to any grade level other than Grade 3 at the district level.

Limitations

Any errors resulting from self-reporting of data or data entry could not be determined. The results of the research apply only to data generated from the NJ ASK 3 Language Arts and Mathematics scores and demographic data from the specific districts sampled in New Jersey.

The study was not an experimental design and therefore cannot determine cause. The samples size for this study was the entire population with at least 25 students enrolled in third grade Therefore estimates about specific characteristics of all New Jersey school districts can be made with a high degree of reliability; there is a low probability, near zero, the results were chance.

The data gathered for this study represents one point in time. The dependent variable data of school district NJ ASK 3 Language Arts and Mathematics proficiency scores was taken from the year 2009. For purposes of this study it is assumed these assessments accurately measured student achievement at the district level. It was also assumed all districts complied with testing regulations to ensure test results were valid.

District socioeconomic data was taken from the American Community Survey (ACS) results. The survey produced five-year, three-year and one-year estimates for each category. This study used the data from the five-year estimates because this provided the largest sample size of the three estimates.

District economically disadvantaged data combines free and reduced lunch

programs and treats them as the same. There are significantly different degrees of income levels combined into one percentage.

Definition of Terms

High-Stakes: "Three conditions must be present for a test or testing program to be considered high-stakes: (a) a significant consequence related to individual student's performance, (b) the test results must be the basis for the evaluation of quality and success of school districts, and (c) the test results must be the basis for the evaluation of quality and success of individual teachers" (Tienken & Rodriguez, 2010).

District Factor Group: These groupings of school districts in New Jersey began in 1975. The purpose of these groupings is to allow student performance on state standardized tests to be compared to student performance from communities with comparatively similar socioeconomic status.

Standard Error of Measurement: The Standard Error of Measurement (SEM) is an estimate of the amount of error or lack of precision one must consider when interpreting a test score (Tienken & Rodriguez, 2010).

Predictive Validity: Predictive validity is the extent to which a score on a scale or test predicts scores on some criterion measure (Cronbach & Meehl, 1955).

Adequate Yearly Progress (AYP): NCLB established the goal of one hundred percent of achieving proficiency at each grade level in Language Arts and Mathematics. AYP targets are established for the years prior to 2014 to identify which districts are on track to achieve the one hundred percent mark. Districts are required to publish their AYP results annually.

New Jersey Assessment of Knowledge and Skills (NJ ASK): The assessment used by

New Jersey in Grades 3 to 8 to determine if districts are meeting AYP targets in Language Arts and Mathematics. Science is administered in Grade 4 and Grade 8. It is administered during the spring of each school year. It was first administered in the spring of 2004.

No Child Left Behind (NCLB): President George W. Bush signed this legislation into law on January 8, 2002. NCLB mandates that states meet the goal of one hundred percent proficiency for all students by the year 2014.

Chapter Summary

The work of Francis Parker marks the beginning of American Progressive educational theory, which explicitly valued the needs of the individual learner in the education process. Early reactions to Progressive theory were characterized by theorists such as William Bagley, which were in response to a sense of inferiority or "falling behind" about American society. Bagley's theories laid the foundation for Essentialism, which prioritized sameness over the needs of the individual learner. These dueling paradigms have dominated the development of educational policy since the early 1900s. Today, Essentialist theories dominate the landscape of American education policy. A problem exists where Essentialism policies lack a solid research base. Maylone (2002) demonstrated how out-of-school variables explained more than 56% of Michigan school districts' high school achievement scores. In some cases, Maylone (2002) applied out-of-school variables to predict school districts' actual high school test scores. Whereas Essentialist theories value the importance of standardized assessments to measure student achievement, Maylone (2002) suggests these standardized assessments may indeed say little about student achievement as a result of the influence of out-of-school variables at

the high school level. Therefore, a need exists to determine if specific out-of-school variables can explain student achievement in a school district at the elementary school level to better explain the value of standardized assessment data during the earliest stages of student development. A dearth of empirical evidence exists regarding the predictive power of out-of-school variables at the elementary level of school districts. This study added empirical results to the limited body of existing literature.

This study examined three overarching research questions:

1. How much variance in NJ ASK 3 2009 test results in Language Arts and Mathematics is explained by out-of-school socioeconomic variables?

2. How accurately can out-of-school socioeconomic and community-level variables predict a school district's percentage of students scoring Proficient or above on the NJ ASK 3 Language Arts and Mathematics sections?

3. Which community-level variables account for the greatest amount of variance in a school district's percentage of students passing the NJ ASK 3?

Chapter II

REVIEW OF THE LITERATURE

The following literature review examines research and articles pertaining to high-stakes public education standardized assessments and the history of dueling paradigms influencing the development of the American public education system. Literature pertaining to NJ ASK 3 and the impact of district demographic data on student achievement data receives particular attention. The work of critical theorists from varying time periods was included to highlight the broader context driving ongoing policy development. The literature review was organized into the following sections: The Essentialist Paradigm, The Progressive Paradigm, Review of Assessment and High-Stakes Policy Development, Technical Characteristics of Standardized Assessments, Importance of Early Childhood Learning and Intervention, Influence of Demographic Factors on Achievement, and Theoretical Framework.

Overview of Existing Literature

Identifying a dominant paradigm guiding the development of educational policy during a particular time period helps to better explain the contextual environment of the variables a study seeks to examine. Evidence for the dominant paradigm during a given period of time can be found in how policy initiatives define the purpose of the public school system.

The purpose of an American public education system was pondered during the earliest days of America's young government. Thomas Jefferson asserted this purpose was to "develop an intelligent citizenry and to provide educational opportunities that guarantee each individual the chance for optimal development" (Tanner & Tanner, 2007

p.4). During the dominance of the Progressive Era, *The Cardinal Principles of Secondary Education* offered another answer. The report concluded, "Education in a democracy, both within and without the school, should develop in each individual the knowledge, interests, ideals, habits, and powers whereby he will find his place and use that place to shape both himself and society toward ever nobler ends" (National Education Association of the United States, Commission on the Reorganization of Secondary Education, 1918, p. 3). This report laid the groundwork for a massive expansion of the public school system by both legitimizing the expectation that every student should graduate from high school and asserting the value of educating the "whole" child. Both Jefferson and *The Cardinal Principles of Secondary Education* articulated a vision of a public school consistent with the tenets of Progressivism. Today the Common Core Standards state that the purpose of public education is to prepare every student to be ready for success in college, a much more limited view of the public school system than Jefferson or *The Cardinal Principles of Secondary Education* recommended and further evidence for the dominance of Essentialist theory. Interestingly, the Eight-Year Study and Thorndike (1924) provide a significant research base for how Progressive strategies have proven to effectively prepare students to achieve the Essentialist vision of the public school. However, present day high-stakes standardized assessment policies, rooted in Essentialist theory, lack a significant research base to demonstrate they will be effective.

Furthermore, a broad body of research has shown that district economic and social demographic factors significantly impact student achievement as measured by standardized assessments (Alspaugh, 1991; Coleman, Hobson, McPartland, Mood,

Weinfield, & York, 1966; Payne & Biddle, 1999; Roscigno & Ainsworth-Darnell, 1999; Sirin, 2005; White, Reynolds, Thomas, Gitzlaff, 1993). The Early Childhood Longitudinal Study Kindergarten Cohort of 1998 conducted by the Department of Education found that the academic skills of economically disadvantaged students are significantly lower than those of their more affluent peers (Fryer & Levitt, 2004). Since *A Nation at Risk* (US Department of Education 1983), education policies increasingly emphasized the need for rigorous standards and assessment as the central strategy to improve student achievement. Over this same period a greater number of states established high-stakes, rewards and consequences, based on data generated from standardized assessments (Amrein & Berliner, December 2002; Amrein & Berliner, March 2002). Presently, reform initiatives aim to further increase the high-stakes attached to standardized assessment data by aggregating the data to determine the degree to which a school district, school and teacher influence student achievement.

Existing literature about the use of high-stakes standardized assessment data to determine the quality and success of a school district can be categorized by two contradictory conclusions. One group of studies concluded that high-stakes accountability systems prove to be an effective reform policy and have improved student achievement over the past two decades (Braun, 2004; Hanushek & Raymond, 2003; Hanushek & Raymond, 2004; Hanushek & Raymond, 2005). These studies also concluded that these policies need to be improved in several areas. Another group of studies concluded that high-stakes accountability systems have done little to improve student achievement and in some cases have done harm (Amrein & Berliner, December 2002; Amrein & Berliner, March 2002; Dorn, 1998; Elmore, 2004; Linn, 2000; Mehrens,

1998; Rustique-Forrester, 2005). There is general agreement within the research that high-stakes accountability systems have not impacted the achievement gap between economically disadvantaged students and their more affluent peers (Hanushek, Raymond, & Rivkin, 2004). There is also evidence to suggest this achievement gap was significantly reduced during the 1970s and 1980s before increasing again during the 1990s (Lee, 2002).

Significance of Existing Literature

The need to identify quality and success in school districts remains great. Harvard Professor and policy advisor Richard Elmore writes, "Considering the magnitude of the task posed by standards-based reform for local school districts and schools, there is shockingly little research and documentation of institutional design and practice in exceptionally high-performing school districts…the knowledge base on which to base advice to local districts on the design of large-scale improvement processes is very narrow" (Elmore, 2004). A need clearly exists to identify quality and success while controlling for district socioeconomic and demographic factors to determine which school districts are worthy of greater study. Existing literature establishes the influence of district demographics on high-stakes standardized test data and in the case of Maylone (2002) the predictive power of this data. Therefore, if out-of-school variables are found to explain significant variance, and in some cases predictive power, in district test scores, as the existing literature suggests, the value of using district test scores to measure the quality of in-school variables may be in question.

Literature Search Procedures

Literature to be reviewed for this study was accessed via online databases:

JSTOR, EPAA, EBSCOhost, ProQuest, ERIC and Google Scholar. A variety of non-experimental studies comprise the majority of studies reviewed. The field of education doesn't produce a large amount of randomized experimental studies. Therefore non-experimental research studies provide the most data about educational practices and serve a critical role within the field of educational research. "The review of related literature involves the systematic identification, location, and analysis of documents containing information related to the research problem…These documents can include articles, abstracts, reviews, monographs, dissertations, books, other research reports, and electronic media" (Gay, Mills, & Airasian, 2009 p.80). The framework outlined in Boote and Beile (2005) for effectively organizing a literature review was applied in this study.

Methodological Issues in Studies of Predictors on Student Achievement

The review of literature pertaining to the variables examined in this study revealed two significant methodological issues. Overall, the body of research relating to predictors of student achievement lacks experimental studies and therefore the literature is limited in its ability to identify reliable casual relationships between variables. The prevalence of conflicting interpretations of similar data oftentimes led to conflicting conclusions and findings. This issue was most often present in the longitudinal studies about the impact of high-stakes assessment policies on NEAP scores at the state level.

Inclusion and Exclusion Criteria for the Literature Review

Studies that met the following criteria were included in this review:

1. Peer-reviewed dissertations or government reports

2. Reported at least statistical significance of findings

3. Published within the last thirty years unless a seminal piece relevant to review of a specific time period

4. Used an experimental, quasi-experimental, non-experimental with control groups, or quantitative empirical study design

The review begins by outlining the evolution of two dueling paradigms of thought about American public education over the past century. Theorists were given particular attention for the extent of their influence on later theorists and the volume of their writing. These theorists/researchers include William Bagley, Edward Thorndike, Frederick Taylor, Francis W. Parker, and John Dewey. Next, a review of the two landmark studies conducted in America about the influences of out-of-school variables on student achievement provides insight about the potential influence of district socio-demographic data on district student achievement. A more thorough examination of the literature pertaining to the variables of household income, parental education, family poverty and lone-parent households provides additional insight into their potential influence on student achievement. Particular attention is given to Maylone (2002) and Jones (2008). Amrein and Berliner (2002a and 2002b) provide the context for examining the debate about the influence of high-stakes testing policy on student achievement on NAEP scores at the state level. Next, the technical characteristics of the NJ ASK are examined and Messick (1995) is reviewed to explain issues of validity and reliability relevant to standardized assessments. Before outlining the theoretical framework of the study, literature pertaining to the effectiveness of early interventions on child development is examined.

Review of Literature Topics

The Essentialist Paradigm

In essence, Essentialist theory about education grew from the clash of two giants (Pauli, 1960). Studying side-by-side with John Dewey and Edward Thorndike, William Bagley became a forceful advocate for transforming the work of education into a profession rooted in the science of educational theory. In his early work, Bagley attempted to establish educational theory by using the new science of psychology (Pauli, 1960) as its foundation. During his time, psychology as a science was also in its infancy and viewed by many as the science of the mind. *The Educative Process* (1907) became Bagley's most comprehensive attempt to define a science of education in which he aimed to reduce education to its simplest terms. He wrote, "In the study of the concrete problems of education we need a guiding principle; we need a formula that will convey every case that is presented; we need to know what education means in its simplest terms" (Bagley, 1907 p. 3). This sample from his early writing foreshadows the Essentialist vision he described in later work.

Acknowledging the virtue in the research of Dewey and Thorndike, Bagley sought "unifying concepts" (Pauli, 1960) between the work of both within education to build a science for educational theory. Bagley's thinking evolved further with *Educational Values* (1912), in which he attempted "to organize the methods of teaching upon a rational basis (Bagley, 1912), but to no avail. During a 1914 debate with David Sneeden, Bagley revealed in strikingly frank terms that he no longer believed a theory for the science of education could be built on the foundation of psychology (Johanningmeier, 1969).

It seems this revelation played a role in furthering Bagley's evolving Essentialist vision. "In his 1918 address to the Normal Department of the NEA, Bagley urged that the traditional distinction between academic subjects and professional subjects be abolished and identified three current modes of preparation which adhered to that distinction" (Johanningmeier, 1969 p. 17). At this point in his career, Bagley begins to reject the study of educational theory and psychology in teacher training programs. Bagley's disillusionment with defining a science for educational theory manifests itself with his final major work, *Education and Emergent Man* (Bagley, 1934). Rather than offering educational theory based on the science of psychology, Bagley describes educational theories as a philosophy. A career born by the desire to define the science of education concluded with a philosophy about the essential elements of the education process.

Bagley (Bagley, 1938) would become the manifesto for Essentialist theory. It was not, however, well received. Bagley noted the very formation of The Essentialist Committee for the Advancement of American Education in February of 1938 attracted criticism before the committee even met (Bagley, 1939). In defense, Bagley rationalized the significance of Essentialist theory as a reaction to "certain incontestable weaknesses in American education. These weakness were traced to the vast upward expansion of the universal school"(Bagley, 1939, p. 329). Bagley interpreted the mass expansion of the public school during the first half of the twentieth century "as the primary causal factor in the relative weakness of American education" (Bagley, 1939, p. 330). From the Essentialist perspective, the mass expansion of the public school created an increasingly heterogeneous student body of high school graduates and therefore a problem for colleges

and universities in their admissions process. Important to note, Bagley doesn't oppose the expansion of the public school; rather, he suggests this expansion caused a relaxation of rigorous requirements to the detriment of the country. Bagley justifies his claim that America is falling behind other developed nations by citing the lack of Americans proportionally represented by Nobel-Prize awards and America's per-capita consumption of literature. Bagley also reports "among all the countries that have embraced the ideal of universal elementary education, ours is apparently the only one in which the expansion of the universal school has not been paralleled by decreasing ratios of serious crime" (Bagley, 1939, p. 333). Consequently, in 1938 Essentialists charge that American education is "appallingly weak and ineffective" (Bagley, 1939, p. 333).

While this charge was roundly rejected by popular Progressive theorists of the time, it is consistent with how present day education reformers justify an increasingly Essentialist platform for public schools. From its earliest roots to the present day, Essentialism exists as a reaction to perceived weaknesses in the greater American society the causes of which are assigned to the American public school system.

After his death, Bagley's theories continued to attract attention. Another influential theorist, Arthur Bestor, advocated Essentialist principles in *The Restoration of Learning* (Bestor, 1955). Bestor expanded on Bagley (1938) by stating that the mission of school is the development of intellectual disciplines through the training of the mind. Bestor suggested the process required to achieve this vision must include rigorous testing and examination. As a result of this rigorous assessment, Bestor suggested students should be ranked and ordered accordingly. He articulated a vision in which the purpose of schooling was to sort and rank students rather than empower students through

experience. Bestor promulgated Bagley's original ideas about education, but Essentialist theory remained widely unpopular through the 1960s and 1970s.

This unpopularity began to shift in 1982. Mortimer Adler's manifesto for education (Adler, 1982) further developed Essentialist theory by describing a single-track Essentialist curriculum for all grades. In Adler's vision of a public school, the only elective class offered would be a choice of one foreign language. Adler's theories viewed the mind as a muscle and rationalized that it could be strengthened through "three distinct modes of teaching and learning" (Tanner & Tanner, 2007 p. 308). These modes included lecture, recitation to organized knowledge, coaching and drilling exercises to develop skill and Socratic questioning and response to deepen understanding about concepts. Adler (1982) provided Essentialist theory with a tantalizing, detailed prescription for American education, and it would greatly influence a soon to be influential national report.

In 1983 three substantial national reports about education called for:

> "curriculum priority in the sciences and mathematics through federal funding, raising the requirements in the sciences and mathematics for college entrance, giving increased emphasis to modern foreign languages and computer literacy, using national standardized tests to clarify the transition through levels of schooling, raising graduation requirements in the four academic "basics" and adding the "new basic" of computer science, increasing academic learning time and reducing electives, eliminating "nonessential" subjects, instituting ability grouping and special programs for the academically talented, increasing the amount of homework and establishing a longer school year and school day,

establishing a federal mechanism for the assessment of student achievement for national, state, and local evaluation and comparison and raising the standards for teaching" (Tanner and Tanner p. 306, 2007).

The three reports, *Educating Americans for the 21st Century* (Coleman, Selby, Cecily, Cannon, et al., 1983), *Action for Excellence* (Task Force on Education for Economic Growth, 1983), and *A Nation at Risk* (National Commission on Excellence in Education, 1983), in totality represent a watershed moment in educational theory. After their publication, Essentialist theory gained increasing dominance over evolving educational policy in the development of the American public education system.

Of the three reports, *A Nation at Risk* (1983) received the greatest attention. Its lead author, Theodore Sizer, not coincidentally also served on Mortimer Adler's Paideia Group, and the influence of Adler's work on *A Nation at Risk* (1983) is unmistakable. Sizer followed his work on *A Nation at Risk* (1983) with *Horace's Compromise* (Sizer, 1984). In totality, Sizer attacked comprehensive high schools and advocated forcefully that schools must adhere to a "less is more" principle and the development of intellectual habits in all students. By characterizing comprehensive high schools as "shopping malls," Sizer created a powerful and lasting metaphor about what schools should not be. Coupled with the alarmingly titled *A Nation at Risk* (1983), a sense of urgency to implement Essentialist theories in public schools continued to grow stronger and remains the dominant educational philosophy today.

Where the *Cardinal Principles of Education* (1918) rationalize a comprehensive American public education for all students, Bagley (1938) theorized public education should focus on a specific program of study for all students. Lost in much of the

backlash to Bagley's initial Essentialist theory was his acceptance of the basic tenets associated with child-centered Progressivism as acceptable methodology to achieve an Essentialist vision. Indeed, Bagley acknowledged many positive attributes of Progressive theory while calling for knowledge to be more clearly organized by subject matter and rigorous standards for learning specific knowledge. Later Essentialists (Adler, 1982; Bestor, 1955; Hirsch, 1987; Hirsch, 1996; Sizer, 1984) have not articulated a balanced view of Essentialist and Progressive theory. Over time, Essentialist theory has increasingly advanced theories in which student-centered Progressive educational theory cannot and does not coexist. Present day reform initiatives like RTTP and Common Core represent an unprecedented increased influence of the federal government in American public education. They also realize that Bagley's vision for the public school minus the Progressive theories he acknowledged had value. Indeed, Bagley was an early voice calling for a greater Federal role in the public school system (W. C. Bagley, 1919) by characterizing American society as "failing" and American education as a national problem impacting all. Bagley did not, however, view Essentialist and Progressive policies as mutually exclusive in the way modern education policymakers appear to.

The Progressive Paradigm

Francis W. Parker offered "a rounded educational doctrine for the study and criticism of teachers" (Parker, 1894) with the publication *of Talks on Pedagogies: An Outline on the Theory of Concentration.* Considered by John Dewey the "father" of Progressivism, Parker served with much distinction as the superintendent of schools in Quincy, Massachusetts and concluded his work by founding the Cook County Normal School. Parker (1894) presented the problem of education succinctly: "In the beginning

of these discussions, the question of all questions and indeed the everlasting question: What is the being to be educated? What is the child? What is the lump of flesh, breathing life and singing the song of immortality?" With stark clarity, Parker established the fundamental principle of Progressivism: child-centered education.

While Parker began defining foundational principles of Progressivism, the American public education system was about to explode. Between 1900 and 1920 enrollment in secondary high schools grew from 500,000 to 2,000,000 students. During the same time span, the number of high schools in America more than doubled from 6,000 to 14,000. Within the context of this massive growth and expansion of the public school system, Progressivism was born. Whereas Essentialism grew out of perceived "weaknesses" of the public education system, Progressivism reacted to the dominant social movements of the early 1900s: social efficiency and scientific management. The massive expansion of the public education system exemplified the values of these social movements. Frederick Taylor (1916), an industrial engineer, outlined his theories about efficiency, management and productivity in *The Principles of Scientific Management* (Taylor, 1911). Taylor advocated the widespread adoption of his management practices to improve democracy and overall social efficiency. Similar to how the detail and clarity of Adler (1982) exemplified Essentialist values, Taylor's *The Principles of Scientific Management* offered the same tantalizing guarantees for an education system desperate to expand.

Convinced that there is "one best way" to complete a task, Taylor advocated time-motion studies be conducted in three steps: (1) the basic movements in a job were catalogued, (2) baseline data for how much time each movement takes was recorded and

(3) a "standard time" or "quickest time" was experimentally determined by recording the time it took one of the fastest workers (first-rate) to complete the task under optimal conditions. This was the process Taylor described to determine the "one-best-way." Taylor viewed management as an exact science without differentiation for the individual worker. He strove to find in everything the "one best way." Another aspect of Taylor's principles was the value placed on external rewards and punishments to motivate the worker. Applied to education theory, Taylor's principles aligned nicely with Bagley twenty years before he would publish his Essentialist manifesto.

Consistent with Taylor, Dewey also valued science in education. In *Democracy and Education* (1916), Dewey put forth his most comprehensive manifesto about educational philosophy, and science is valued throughout. In doing so, Dewey concedes the value present in the division of labor and standardized practice in the workplace but questions the lack of intrinsic rewards generated by the work. "The tendency to reduce such things as efficiency of activity and scientific management to purely technical externals is evidence of the one-sided stimulation of thought given to those in control of industry–who supply its aims" (Dewey, 1916). For Dewey, the "aims" of education versus the "aims" of Taylor's *The Principles of Scientific Management* were vastly different. He viewed Taylor's "aims" as "fixed and rigid; it is not a stimulus to intelligence in the given situation, but is an externally dictated order to do such and such things (Dewey, 1916)." Dewey conceived the purpose of education to be lifelong learning with no end in mind, no "aim." Dewey presents this view as a more balanced view of learning by allowing for the individual needs of the learner to be accounted for. Dewey's *Democracy and Education* (1916) in some ways exists as a critique of Taylor's

The Principles of Scientific Management. Where Taylor called for "one best way," Dewey recommends a plurality of methods to achieve optimal growth in both industry and the worker.

Dewey's *Democracy and Education* (1916) is a foundational expansive document about Progressivist thought in education that has stood the test of time. Dewey (1916) ties educational theory to business, as does Taylor, and to politics by equating his views about education with effective democracy. The business references serve as both a critique of Taylor's theories and a cautionary tale about the aspirations of business driving educational theory. "Skill obtained apart from thinking is not connected with any sense of the purposes for which it is to be used. It consequently leaves a man at the mercy of his routine habits and of the authoritative control of others, who know what they are about and who are not especially scrupulous as to their means of achievement" (Dewey, 1916 p.110). As for democracy, Dewey rationalizes the aspirations of public education and Democracy to be inseparable.

Dewey concludes that the ultimate purpose of education must be to teach the student to think. "The sole direct path to enduring improvement in the methods of instruction and learning consists in centering upon the conditions which exact, promote, and test thinking…the important thing to bear in mind is that thinking is the method, the method of intelligent experience in the course which it takes" (Dewey,1916) By tying the act of thinking to method, Dewey lays the foundation for an evolving education philosophy to value the process equally with the product. This line of thinking will give birth to an authentic assessment movement and provide rationale for the need to teach the "whole" child. This one aspect of Progressivism stands in particularly stark conflict with

Essentialist theory. Dewey (1916) theorizes a strong correlation between a viable democracy and a public education experience where students develop a method of thinking resulting from intrinsic inquiry. With this comprehensive vision for public education articulated fully in Dewey (1916), a Progressive Era would soon become increasingly popular.

The peak of the Progressive Era produced its most influential study. By 1930 the number of students attending high school had grown from one million in 1900 to over ten million. This rapid growth, along with questions about how high schools could better serve all students, led to the convening of educational leaders at a 1930 conference in Washington, D.C. At this conference, a glaring obstacle to meaningful reform of the high school became increasingly clear. It was acknowledged that the requirements colleges placed on admissions were driving the courses of study offered by high schools. The admission requirements were rooted in and justified by the theories of mental discipline, the belief that one particular course of study was superior to other courses of study to prepare students for success in college.

Recognizing that this assumption had never been tested, the Progressive Education Association commissioned a study to examine the relationship of the high school experience to college success. The result was the Eight-Year Study, which examined 1475 pairs of students attending college between 1936 and 1939. The Progressive nature of the study, however, took place in 1932 when 30 high schools (private and public) were chosen to participate in the study. These high schools were identified based on location, demographics, and desire for reform. The study intentionally chose schools that expressed eagerness to reform their curriculum, school

design, and competency in order to be successful in this endeavor. The schools began their work during the fall of 1933. At first, the task appeared overwhelming and produced an often-cited quote from a participating school principal. "My teacher and I do not know what to do with this freedom. It challenges and frightens us. I fear we have come to love our chains" (Aiken, 1942, p.16). This paralysis did not last.

Each high school was given a high degree of autonomy and independence to make changes to curriculum and school design. The curriculum process was not standardized and became highly experimental in each school. The schools did develop common principles to guide their evolution and school leadership became highly democratic. The first cohort of graduates entered college in 1936. At this point, 25 of the 179 colleges approved for students from these experimental high schools to attend were selected for close examination. The 25 schools were selected based on the abundance of students from the experimental high schools accepted to the college. For these high school graduates, college admissions were based on a detailed principal recommendation and historical narrative of the student's school experience, which included diagnostic data about aptitude and achievement. Admission was not based on criteria for a prescribed course of study.

The findings of the Eight-Year Study were published during the summer of 1940. After analysis of 18 different outcomes, the study demonstrated that "the graduates from the most experimental schools were strikingly more successful than their matches' (Aiken, 1942, p. 113). The "matches" referred to students that had graduated from traditional high schools without an experimental curriculum but a highly prescribed course of study. Further analysis demonstrated that students from the most experimental

high schools were the most successful in college. "The results of this study seem to indicate that the pattern of preparatory school program which concentrates on a preparation for a fixed set of entrance examinations is not the only satisfactory means of fitting a boy or girl for making the most out of the college experience. It looks as if the stimulus and the initiative, which the less conventional approach to secondary school education afford sends onto college afford better human material than we have obtained in the past" (Aiken, 1942). Greater justification for the value of Progressive education has yet to be produced.

Interestingly, Edward Thorndike provided Progressive education with further evidence of the fallacy of mental discipline driving college admissions rationale to require a prescribed course of study to earn admissions. Thorndike's influence on the evolution of America's public education system cannot be overstated. While Frederick Taylor's theory of scientific management is often cited as explanation for the design of the American public education system. Thorndike's influence was equally great. His seminal work, *Animal Intelligence: Experimental Studies* (Aiken, 1942), gave birth to the Situation and Response Theory of learning. Thorndike did not value the role of reasoning in the learning process and theorized that the results of his experiments with animals demonstrated that learning takes place when students associate the correct actions with successful responses. Thorndike was convinced this model could "explain all aspects of learning… higher animals, including man… manifest no behavior beyond exception from the laws of instinct, exercise, and effect…learning is connecting….The mind is man's connection system. Purposes are as mechanical in their nature as anything else" (Tomlinson, 1997 p. 369).

With a mechanical view of learning as his foundation, it is ironic that Thorndike's work eventually demolished the theory of mental discipline. "In two celebrated experiments, Thorndike showed there is little or no transfer of learning between domain specific tasks and that no subject is more effective than any other in developing a child's intelligence. The classics have no special value in disciplining the intellect, and a general education, in contrast to Charles Eliot's famous claim, was not the best preparation for life" (Tomlinson, 1997, p. 372). Thorndike interpreted these results to suggest intelligence is hereditary and certain gene pools created superior minds. Dewey and Progressive educators interpreted these results much differently. For them, Thorndike's work justified the value of experimental curricula and its effectiveness was further proven by the results of The Eight-Year Study.

Review of Assessment and High-Stakes Policy Development

Edward Thorndike's research expanded beyond the influence of theories about learning. Some of his most consequential work related to the emerging field of assessment and intelligence testing. While testing in schools can be traced back to the days of the one-classroom schoolhouse and teacher-made tests, the twentieth century saw the invention of the standardized assessment. America's engagement in World War I prompted the army to devise a system to screen the aptitude of recruits for leadership responsibilities. As a result, The Committee on Classification of Personnel from 1917 to 1919 was commissioned by the government to conduct the first mass effort to measure intelligence. Thorndike worked as a member of this committee. The committee created a "Beta Form" for illiterate recruits and an "Alpha Form" for literate recruits to be administered to 2 million soldiers by 1919. Not long after these standardized assessment

were given to army recruits, school children were administered intelligence testing on a national scale. Since these initial assessments, standardized assessment of student learning has continued to play an increasing role in America's public education system.

For much of the twentieth century standardized assessments were used to determine student proficiency levels at the district level and achievement levels in specific subject matter at the individual school level. The use of this data began to shift with the passage of the National Defense Education Act (NDEA) (Federal Government, 1958) in 1958. In response to Russia's launch of Sputnik, "NDEA focused attention on increasing student achievement in science, mathematics, and foreign languages. The act allowed science scholars to assist in constructing a science curriculum in high schools, while reaffirming the rights of local governments to control their school curriculum, administration, instruction, and personnel" (Jones, 2008). NDEA would be a prelude to unprecedented questions from the federal government about student outcomes. Before the 1960s, schools were not seen as responsible for proving student outcomes. "While a good school prior to 1965 provided students and teachers with the materials associated with education, The Elementary and Secondary Education Act (ESEA) shifted accountability towards a student-centered model. The government declared that high-quality schools produced favorable student outcomes" (Cuban, 1993).

The Elementary and Secondary Education Act of 1965 proved to be a watershed moment for the American public education system by establishing a framework for increased accountability of schools regarding specific student outcomes. This foundational principle of increased accountability established by the federal government in 1965 eventually led to the growth of standards-based assessments and the development

of high-stakes education policy associated with quantifiable student outcomes. Initially these assessments and high-stakes policy mostly impacted students and schools at the high school level.

The first round of national standards based assessments during the 1970s were consistent with a "back-to-basics" philosophy and focused on basic skills and minimum competency exams (Jones, 2008). These assessments provided students, families, and schools with individual student performance data and, in some cases, data about national norms for individual students. Initially, there were no high-stakes attached to this first wave of standards-based assessment. This would not last.

Three years after the National Commission on Education released *A Nation at Risk* (1983), 41 states raised high school graduation requirements, 33 states initiated student competency tests, and thirty states required teacher competency tests. This relatively small and unscholarly document with the alarming title effectively tied the improvement of student learning to higher standards and more tests. This idea was not new and continues to go largely unchallenged. Indeed, this idea was similar to the ideology put forth by Essentialists for decades. There was, however, a new idea presented in *A Nation at Risk* (1983). "The commission recommended that states institute higher standards and administer assessments to hold schools accountable for meeting those standards. These assessments became known as high-stakes tests." The idea that schools would be evaluated based solely on students" standardized test scores had not yet been a part of education policy. Since *A Nation at Risk* (1983), high-stakes standardized assessments data has dominated education public policy while the research about the effectiveness of this policy remains inconclusive at best.

Goals 2000: Educate America (P.L.103-227) (Federal Government, 1994) was signed into law by President Bill Clinton on March 31, 1994. The law provided $100 million to states for the development of school improvement plans, sub grants to schools, and the creation of grant awards for professional development. More importantly, the legislation established eight goals and created a framework for measuring student outcomes. It also led to the development of national standards with the formation of a National Education Standards and Improvement Council. The law required states receiving federal funds to set up systems of standards and assessments to measure student outcomes. While Goals 2000: Educate America would be defunded prior to the enactment of NCLB, its passage demonstrated how both Democratic and Republican political parties had become aligned with th Essentialist views of American public education. In New Jersey, this federal legislation led to the creation of New Jersey Core Curriculum Content Standards (NJCCCS) in 1996 and the implementation of statewide testing in Grades 4, 8 and 11.

"The reauthorization of the 1965 Elementary and Secondary Education Act (ESEA, P.L. 89-10), known as the No Child Left Behind Act (NCLB P.L. 107-110), in 2002 cemented test-based policymaking into the education landscape during the first decade of the new millennium" (Tienken, 2011). NCLB established a goal for 100% of students to demonstrate proficiency in Language Arts and Mathematics as measured by statewide standardized assessments by the year 2014. Furthermore, NCLB required all states receiving Title I money to disaggregate student test data by subgroups and then make progress toward eliminating achievement gaps across all subgroups. NCLB also included sanctions for failure to demonstrate Adequate Yearly Progress (AYP) on state-

wide standardized assessments. Prior to NCLB, the high-stakes association with standardized assessment data often came in the form of results being published in the local community. NCLB established more rigorous sanctions for schools failing to make AYP, including the potential elimination of federal funding. Since the inception of NCLB, studies have tried to determine its effectiveness.

One group of studies finds high-stakes assessment polices and practice improve student learning (Braun, 2004; Carnoy & Loeb, 2002; Center on Education Policy, December 2010; Center on Education Policy, October 2009; Center on Education Policy, September 2010; Hanushek & Raymond 2004). A second group of studies finds high-stakes assessment polices to be harmful to student learning and, in some cases, to lead to widening of the achievement gap between racial groups (Amrein & Berliner, December 2002; Amrein & Berliner, March 2002; Nagaoka, & Roderick, 2005; Rustique-Forrester, 2005).

NAEP vs. State Assessments.

President of Harvard and founder of the Education Commission of the States (ECS), James Bryant Conant's *Shaping Educational Policy* (Conant, 1964) called for "counterbalance" at the state level to the growing influence of federal policies on the education system. ECS remains in existence today, governed by state governors. Its most influential contribution remains the creation and administration of National Assessment of Educational Progress (NAEP). The first national NAEP assessment was administered in 1969 and continues to be administered every two years.

The NAEP is often referred to as "the nation's report card." It assesses student achievement at the national level in Grades 4, 8 and 12. One version of the NAEP, also

overseen by the U.S. Department of Education, is administered every two years in reading and math for students in Grades 4 and 8 and trends back to the 1990s. The long-term version of the NAEP is given every four years and trends back to the 1970s. NAEP differs from state level tests in several important areas:

- The NAEP assesses samples of students, whereas state assessments must test all students. Also of note, each NAEP participant takes only a portion of the larger assessment, not the entire assessment.
- The NAEP includes different content and design compared to state tests. Independent test proctors, not teachers familiar with the students, administer it.
- The NAEP is not designed to align with particular state standards. The NAEP is aligned with a framework developed by a National Assessment Governing Board appointed by the Secretary of Education.
- The NAEP proficiency and state level proficiency definitions are often not similar. While state level proficiency rates vary from state to state, most are less ambitious than NAEP cut-off levels.
- NAEP scores are not tied to high-stakes for individual students, teachers, schools, or districts.

Some important limitations of NAEP are worthy of note:

- Student motivation on the NAEP may not be optimal because students' scores are not reported and are not tied to specific rewards or consequences.

- NAEP may not assess what students are taught in the classroom since it is not aligned with one particular set of state standards. Therefore, teachers are unlikely to tailor instruction to NAEP content. (Center on Education Policy, September, 2010)

Hanushek and Raymond (2004) examined the growth of student performance between fourth and eighth grade on NAEP scores between states from the early 1990s to 2004. Using a predictive model of analysis, Hanushek and Raymond (2004) "find that the introduction of accountability systems into a state tends to lead to larger achievement growth than would have occurred without accountability. The analysis, however, indicates that just reporting results has minimal impact on student performance and that the force of accountability comes from attaching consequences such as monetary awards or takeover threats to school performance. This finding supports the contested provisions of NCLB that impose sanctions on failing schools" (Hanushek & Raymond, 2004, p. 33). Also of interest, Hanushek and Raymond (2004) found that differences in NAEP scores across states couldn't be explained by differences in spending. Furthermore, while the evidence of the achievement gap persisted, Hispanics were found to benefit more than African Americans from statewide accountability systems. Hanushek and Raymond conclude, "We find consistent evidence that introduction of state accountability had a positive impact on student performance during the 1990s. Specifically, states that introduced consequential accountability systems early, tended to show more rapid gains in NAEP performance, holding other inputs and policies constant. This is consistent with our prior estimates of the effects of accountability for aggregations of all students in each state" (Hanushek & Raymond, 2003). The state accountability systems diverge

considerably in the types of consequences attached to performance (Hanushek & Raymond, 2004, p.19).

A second group of studies found less a favorable impact of high-stakes testing polices on student achievement (Amrein & Berliner, March 2002; Jones, 2008; Maylone, 2002; Nagaoka & Roderick 2005; Rustique-Forrester, 2005; Tienken, & Rodriguez 2010). Two large-scale studies published by the same authors in 2002 (Amrein & Berliner, March, 2002; Amrein & Berliner, December, 2002) marked the first national studies of test data and high-stakes testing polices. Amrein and Berliner (2002a) examined 18 states identified as having high-stakes testing programs and tried to determine if high-stakes testing policies were affecting student learning. Interestingly, rather than accepting the premise of student learning being defined by statewide standardized tests, Amrein & Berliner (2002a) asked if evidence for transference of student learning is demonstrated by increased achievement on more commonly used standardized assessments: ACT, SAT, NAEP and AP tests. Braun (2004) and Hanushek and Raymond (2004) question the validity of Amrein and Berliner (2002a & 2002b) with contradictory findings. Neither Braun (2004) nor Hanushek and Raymond (2004) question Amrein and Berliner's (2002a) assertion that "high school graduation exams affect students from racial minority backgrounds in greater proportions than they do White students. If these high-stakes tests are discovered not to have their intended effects, if they do not promote the kinds of transfer for learning and education the nation desires, the mistake will have greater consequence for America's children of color" (Amrein & Berliner, 2002, p .8). Braun (2004) and Hanushek and Raymond (2004) both

acknowledge flaws in high-stakes testing policies but neither contemplates the issues of social justice raised by Amrein and Berliner (2002a & 2002b).

Amrein and Berliner (2002a) also question the validity of state developed high-stakes assessment, noting: "The National Research Council cautions that 'An assessment should provide representative coverage of the content and process of the domain being tested, so that the score is a valid measure of the student's knowledge of the broader [domain], not just the particular sample of items on the test" (Amrein & Berliner, March, 2002, p. 15). Because high-stakes tests include fewer items than needed to assess a domain accurately and testing time is rarely extensive enough, these standardized assessment include a thread of invalidity. Braun (2004) and Hunsheck and Raymond (2004) also fail to address the questions of validity raised by Amrein & Berliner (2002a).

Similar to Hanushek and Raymond (2004), Braun (2004) reviewed the performance of states on the NAEP mathematics assessments in Grades 4 and 8 between 1992 and 2000. Braun (2004) compared the performance of states labeled as high-stakes testing states by Amrein & Berliner (2002b) to other states also participating in the NAEP but lacking high-stakes testing policies.

Amrein and Berliner (2002a) identify 18 states with high-stakes testing policies. They analyzed the achievement of students from these states on four assessments: SAT, ACT, NAEP and Advanced Placement exams. Amrein and Berliner (2002a) suggest the effectiveness of high-stakes testing policies are best determined by the evidence of transfer between improved student achievement on state-wide standardized assessment and the previously cited external measures. All of the states examined by Amrein and Berliner (2002a) had established a mandatory high school graduation exam plus other

varying degrees of high-stakes policies. Specific to NAEP scores, Amrein and Berliner (2002a) analyzed "state gain" by determining if each state's NAEP scores outperformed national scores. It is important to note that "the change for the state could be positive but just not as large as the nation's" (Braun, 2004) and still by interpreted as negative by Amrein and Berliner (2002a). Using this formula, Amrein and Berliner (2002a) identify 8 states with positive gains, 3 with negative gains, 2 with zero gain, and 5 without available data. While it would appear that this data supports a theory beneficial to high-stakes testing policy, Amrein and Berliner (2002a) explain that the results are not reliable because of an association between "state gain" and the "change in the percent excluded from NAEP over the same time period (r =0.39) undercuts the interpretability of the results" (Braun, 2004, p. 4).

Braun (2004) conducted a re-analysis of Amrein and Berliner (2002a) by expanding on the question in four critical areas:

1. Inclusion of a parallel analysis of 32 other states as well as the original 18
2. Inclusion of more comprehensive measures of states' education reform policy
3. Interpretation of "state gain" as informed by estimated standard errors.
4. Analysis of Grade 8 data for changes from 1992 to 2000 instead of 1990 to 2000 to make the analyses for Grades 4 and 8 more comparable with additional data. (Braun, 2004, p. 4)

Braun (2004) awards credits to states based on the strength of consistent improvement relative to improvement over the nation. This strategy helps limit the influence of outliers in the data. Braun (2004) found "High-stakes states are more likely to show strongly consistent improvement relative to the nation than low-stakes states and

much less likely to show strongly consistent lack of improvement relative to the nation. The story remains qualitatively the same if we compare the groups with less stringent cut-offs. In summary high-stakes testing states that participated in the NAEP mathematics assessments in both 1992 and 2002 typically showed improvement relative to the nation while low-stakes testing states...typically showed lack of improvement relative to the nation" (Braun, 2004, p. 8).

Rationalizing that the intent of most reform efforts is to improve the achievement of the lowest performing students, Braun (2004) analyzed NAEP mathematics student data from the 25th percentile for both Grades 4 and 8. "There is some robustness to our findings inasmuch as the analyses at the 25th percentile produced similar results.... Consequently, our conclusions differ from those in Amrein and Berliner" (2002a).

Braun (2004) is not the only study to take issue with Amrein and Berliner's (2002a) analysis and conclusions. Carnoy and Loeb (2003) rely on a regression model with the state as the unit of analysis of the same time period. "They find a relatively strong positive association between gains and the accountability index in Grade 8, especially for Black and Hispanic students but a much weaker, though still positive, association in Grade 4....Despite substantial methodological differences in the two approaches, the general tenor of the findings in Carnoy and Loeb (2003) with respect to NAEP results is consistent with ours. With the data available, there is no basis for rejecting the inference that the introduction of high-stakes testing for accountability is associated with gains in NAEP mathematics achievement through 1990s. Moreover, the strength of the association between states' gains and a measure of the general

accountability efforts in the states is greater in the eighth grade than in the fourth" (Braun, 2004, p. 33).

Amrein and Berliner (2002b) elaborate more on the history of policy development at the state level and the relationship between gains/losses as measured by NAEP scores over a longer period of time. In this study, 28 states are identified as high-stakes based on their policies associated with standardized test results. The parameters of Amrein and Berliner (2002b) remain consistent with Amrein and Berliner (2002a). Unlike Braun (2004), Raymond and Hanushek (2003) and Carnoy and Loeb (2003), Amrein and Berliner (2002b) conclude "there is no compelling evidence from a set of states with high-stakes testing policies that those policies result in transfer to the broader domains of knowledge and skill for which high-stakes test scores must be indicators' (Amrein & Berliner, 2002b).

The Center on Education Policy (CEP) October 2009 Report titled *State Test Score Trends 2007-2008, Part 3: Are Achievement Gaps Closing and Is Achievement Rising for All?* conducted three types of analysis: Grade 4 state test results at three achievement levels, gaps between subgroups in the percentage of students scores at each of the three achievement levels, and gaps between subgroups in average test scores at the three grade levels. The CEP October 2009 Report found all subgroups showed more gains than declines in grade 4 at all three achievement levels. In most states it was determined that gaps have narrowed between subgroups and most often because the lower-performing subgroup performance went up. The main unit of analysis was the trend line between 2002 through 2008. However, CEP (October 2009) found, "Despite the progress being made, the nation still has a long way to go to close the achievement

gaps between student subgroups" (Center on Education Policy, October 2009, p. 17). This macro-level study does not account for the large discrepancy between state standards and varying definitions for proficiency from state to state. Overall, the study demonstrates how states have made progress in closing the achievement gap but the progress is "slow and uneven" (Center on Education Policy, October 2009, p. 22).

State Test Score Trends Through 2008-2009, Part 1: Rising Scores on State Tests and NAEP (Center on Education Policy, September 2010) studied whether state-level trends in NAEP reading and mathematics results are consistent with state test scores. The study included 23 states with comparable data from 2005 through 2009. Trends were compared at Grades 4 and 8 with students achieving proficiency on state tests compared to achieving basic levels on the NAEP. The study identified three general findings:

- Since 2005, test scores increased in most states with sufficient data. Of the 21 states studied in Grade 8 reading, 20 showed gains in the percentage reaching the proficient level on their state test, and 17 showed gains in the percentage reaching the basic level on the NAEP. Of the 18 states with mean score data, 15 showed mean score gains on their state tests in Grade 8 reading, and 15 exhibited mean score gains on NAEP.
- Within the same state, trends on NAEP usually moved in the same direction as trends on the state tests.
- Gains on state tests tended to be larger in size than gains on NAEP (Center on Education Policy, September 2010, p. 10).

The study included two more years of data than previously reported to extend an early study's trend lines. "As a final analysis, we sought to determine whether there was

a correlation between the size of the gains on state tests and the size of the gains on the NAEP by computing statistics called correlation coefficients. In other words, was there evidence to suggest that the larger the gain a state made on its state test, the larger the gain it made on NAEP?" (Center on Education Policy, September 2010, p. 15). Between the periods of 2005 to 2009, correlations were weak to moderate. Little relationship was reported between the size of a gain or decline and performance on the NAEP. The study concluded by offering some explanation for the larger gains on NAEP versus state assessments, including instruction being more closely aligned to state standards and score inflation on state tests due to the high-stakes nature of the results.

Similar results were reported in *State Test Score Trends Through 2008-09, Part 2: Slow and Uneven Progress in Narrowing Gaps* (Center on Education Policy, December 2010) Unlike CEP, (Chudowsky & Chudowsky, 2010), examined state level assessments to determine the degree to which achievement gaps persist between subgroups of students. While the study notes every major student group has made gains since 2002 on state reading and math tests, achievement gaps have not always narrowed. The achievement gaps between subgroups remain "large and persistent…although gaps have narrowed more rapidly for some groups than for others; at the current rates of progress it would take many years to close most gaps "(Center on Education Policy, December 2010).

Most recently, CEP (Center on Education Policy, April 2011) concluded student achievement at Grade 8 in reading and math has not stagnated. Rather, it has gone up on most state assessments as well as other assessments. The findings included:

- At all three achievement levels, a large majority of states made gains in

reading and math performance at Grade 8.

- In both reading and math, a larger proportion of states showed gains at Grade 8 than at Grade 4 or high school. This was true at all three achievement levels.
- In Grade 8 math, every state with sufficient data made gains in the percentage of students reaching the Advanced level, and all but one of these states showed gains at the Proficient level as well.

The study does identify a widening achievement gap between subgroups at the advanced level on state assessments:

- Achievement gaps at the advanced level widened for all subgroups at Grade 8 in the majority of states with sufficient data.
- In Grade 8 math, the gap between low-income students and students who are not low-income widened in all but one of these states.
- The Grade 8 math gap between the White subgroup and the higher-scoring Asian American subgroup has also widened in the majority of states analyzed.

The most striking conclusion from CEP (Center on Education Policy, April 2011) is evidence that Grade 8 student achievement growth on state assessments and NAEP is equivalent to student achievement levels in Grades 4 and 12. Previously Grade 8 data had been found not to be improving, similar to Grades 4 and 12 achievement data, whereas a RAND 2004 study of Grade 8 student performance on NAEP during the 1990s and the Trends in International Math and Science Study (TIMSS) found that U.S. eighth graders ranked below average while U.S. fourth graders ranked at the national average (Center on Education Policy, April 2011, p. 11).

High-Stakes Testing Policies

Amrein and Berliner (2002a) identify states with high-stakes testing policies based on the degree to which they have adopted 10 high-stakes polices. These policies include high school graduation exam, promotion exam, public report cards, identifying low performing schools, money awards to schools, money awards to staff, threat of closure, threat of replacing staff, school choice, and money awards to students. Amrein and Berliner (2002a) conclude "Although states may demonstrate increases in scores on their own high-stakes tests, transfer of learning is not a typical outcome of their high-stakes testing policy" (Amrein & Berliner, 2002, p. 52).

Promotional requirements, or the end of social promotion, are one example of an increasingly popular high-stakes policy based on standardized tests scores. Roderick and Nagaoka (2005) examined the impact of Chicago Public Schools promotional requirements in the third, sixth and eighth grades on students retained between 1997 and 2000. During this time period, scores on the Iowa Test of Basic Skills (ITBS) determined sixth and eighth grade student promotions from the third grade. "As a result of this policy, Chicago retained from 7,000 to 10,000 students per year in these grades; nearly one in five third graders and one in ten sixth and eighth graders were subject to the policy" (Nagaoka & Roderick 2005, p.309). The study aimed to answer three central questions:

1. Did the extra year of instruction allow retained students to raise their test scores to meet the promotion standards their second time in the same grade?

2. How did retention under high-stakes testing and Chicago's use of multiple chances to pass the promotion test shape retained students' subsequent progress,

including the probability of special education placement, being retained again, or rejoining their age-appropriate classmates?

3. Did retention lead to higher achievement for theses students than if they had been promoted to the next grade? (Nagaoka & Roderick, 2005 p.310)

During the initial years of the retention policy, the Chicago Public Schools maintained strict adherence to the policy with few waivers. Consequently, in 1998 and 1999 most third graders with reading scores just below the cutoff were retained. In 2000, however, the majority of students with these same scores (near the cutoff score) were promoted. Roderick and Nagaoka (2005) analyzed these similar cohorts of students to better determine the effectiveness of retention. Through comparison of the students retained and those promoted with similar third grade scores on the ITBS, Roderick and Nagaoka (2005) concluded retention did not offer significant academic benefits to students after year one of retention. There was evidence the students did experience academic gains during the retention years but those gains were not maintained when compared with the students not retained with similar ITBS scores in third grade. Students that experienced double retention and those retained and then placed in special education did not demonstrate significant academic improvement as a result of retention. Roderick and Nagaoka (2005) conclude: "In the third grade, there is no evidence that retention led to greater achievement growth two years after the promotional grade. In the sixth grade, we find that retention was, in fact, associated with lower achievement growth. Moreover, there is evidence that retaining students under Chicago's promotional policy significantly increased the likelihood of placement in special education" (Nagaoka & Roderick, 2005, p. 331).

Rustique-Forrester (2005) reports findings from a study examining consequences of England's accountability polices between 1998 and 2001, and Olson (Olson, 2004) observed significant similarities between NCLB and England's high-stakes policies aimed at raising standards and student achievement. The Education Reform Act of 1998 in England led to the creation of national standards and assessments aimed at increasing accountability for student achievement at the school level. The policies evidence the Essentialist paradigm view of a public school. Rustique-Forrester (2005) notes in the decade following the implementation of these Essentialist policies, the number of exclusions from schools at the primary and secondary levels increased 400%. A series of studies found school expulsion rates were strongly correlated with student demographic variables, in particular race, class and, gender (Rustique-Forrester, 2005, p. 9). Furthermore, Rustique-Forrester (2005) found accountability policies increased the pressures and incentives to exclude students who are low-performing. The findings offer evidence of unintended consequences that may result from high-stakes testing policies in the United States.

More recently, Lee (2007) examined long-term trends in U.S. national math and reading scores between 1971 and 2004. While small improvements in reading and math are evident, there was no significant change in the trajectory of the trend lines. They remain mostly linear over the 33-year history of the NAEP. The achievement gap trend lines are different. They follow more of a curvilinear pattern, with closing of achievement gaps in the 1970s and early 1980s, followed by an increase in the late 1980s and 1990s with no significant progress since NCLB enactment in 2002 (Mintrop & Sunderman, 2009).

Synthesis.

The research about the impact of high-stakes testing policy on student achievement remains mixed at best. One area in which researchers agree relates to the achievement gap and its persistence even with high-stakes testing policies in place. The most ardent and prolific defenders of high-stakes testing policies, Raymond and Hanushek (2003), attempt to discredit Amrein and Berliner (2002a and b). Pereira (2011) artfully notes, "Although Raymond and Hanushek attempted to disprove Amrein and Berliner (2002a) theories of negative high-stakes influences on student learning, what they did instead was emphasize just how insignificant the point advantages truly were. Acknowledging the NAEP mathematics scores range on a 0 to 500 scale, even at the greatest recorded point advantage of 5.3 for the Grade 4 test during the 1992-2000 period, that is still only a 1.06% increase. Therefore, this "advantage" proves completely insignificant to use as evidence of a positive correlation between increased student achievement and the implementation of high-stakes attachment to tests, as Raymond and Hanushek maintained "rigorous analysis reveals that accountability policies have had a positive impact on test scores during the past decade" (2003, p. 50) (Pereira, 2011, p. 49).

The debate about high-stakes testing policies will continue. Pereira (2011) demonstrates the thin body of research to support the overall effect of these policies. Combined with the lack of research to support the effectiveness of the Essentialist theories these policies are rooted in, skepticism about said policies is warranted. Indeed, the most telling data analysis of NAEP scores may be by Mintrop and Sunderman (Mintrop & Sunderman, 2009). The achievement gap trend lines of NAEP scores follow more of a curvilinear pattern, with closing of achievement gaps in the 1970s and early

1980s followed by an increase in the late 1980s and 1990s with no significant progress since the enactment of NCLB in 2002. A watershed moment for Essentialist theories is marked by the publication of *A Nation at Risk* (National Commission on Excellence in Education, 1983). Can it simply be coincidental that the increased influence of Essentialist thinking over education policy coincided with the reversal of gains made in closing the achievement gap as evidenced by NAEP scores?

Technical Characteristics of NJ ASK 3

Standardized tests generate one of two types of test scores: norm-referenced or criterion-referenced. Norm-referenced scores compare students to other students in a similar peer group. Criterion-referenced scores interpret student scores against a specific set of criteria or standards. Criterion-referenced tests are commonly known as standards-based assessments. The NJ ASK 3 is an example of a criterion-referenced standardized assessment. The NJ ASK 3 was first administered in the spring of 2003 to all students in Grade 3. The NJ ASK 3 includes an assessment in Language Arts Literacy and Mathematics. It is the earliest criterion-referenced assessment administered statewide in New Jersey.

The NJ ASK Grades 3 and 4 Technical Report 2008 states that the intent of NJ ASK assessments is to be an early indication about student progress toward mastery of New Jersey Core Curriculum Content Standards (NJCCCS). NJ ASK 3 produces scale scores for each student with three levels:

100–199	Partially Proficient
200–249	Proficient
250–300	Advanced Proficient

The Test Design for Language Arts Literacy includes two content areas: Reading and Writing. Since 2009, two types of reading passages are included on the NJ ASK 3, narrative and informational, in order to assess two reading clusters (a group of related test questions on a single standard).

1. Working with/Interpreting Text tests the following skills:
    - Recognizing the central idea or theme
    - Recognizing supporting detail
    - Extrapolating information/following directions
    - Paraphrasing/retelling
    - Recognizing text organization
    - Recognizing a purpose for reading
2. Analyzing/Critiquing Text tests the following skills:
    - Questioning, clarifying, predicting
    - Predicting tentative meaning
    - Forming opinions about text and author techniques
    - Making judgments/drawing conclusions
    - Interpreting textual conventions and literacy elements

The Writing content is assessed through two expository prompts: a brief verbal prompt and a poem to introduce a topic. Students are expected to draft a composition as a response to each prompt.

Test items for the NJ ASK include multiple choice, written response, and open-ended (OE). Before 2007, open-ended test items were scored by two separate scorers and the final score was determined by averaging the two scores. Since 2007, open-ended test

items are scored by one scorer. *The NJDOE Technical Manual for NJ ASK 3 and 4 2008* notes how the change from two scorers to one scorer is the equivalent of using a new item. Therefore, a Rescore Equity Study was conducted to determine the impact of this change on the reliability of test scores. The study (New Jersey Department of Education, 2009a) aimed to answer three questions:

1. How will changing scoring from two to one rater affect student scores and score distribution?

2. If "replacement" OE anchor items are used, are the equating results obtained substantially different?

3. How will this shift in OE scoring procedures affect the cut score and the percentage of students in each performance level?

To answer these questions, Education Testing Services conducted a rescore of 2005 and 2006 data. The study found the following:

- State mean and overall score distribution did not change –correlation between scores in all areas was greater than .99
- Scores for individual students did change. Therefore an automatic rescore was enacted for all students with a raw score within two raw points of the Proficient cut score.

The NJDOE NJ ASK 3-8 Technical Manual 2009 noted all Writing responses no longer were scored by two raters, only one. Of note, there is no mention in the Manual of any Rescore Equity Study to determine the impact of this procedural change in scoring on students' scores. A new standard-setting procedure is also introduced in the 2009 Manual. A Bookmark Procedure was completed to determine appropriate cut-off scores

for Proficient and Advanced Proficient levels. Panelists were asked to review a booklet of test items organized from "easiest" to "hardest." Panelists placed two bookmarks in the booklet, one bookmark at the page where they believed a Proficient student had no better than a 2/3 chance of answering the questions correctly and another where they believed an Advanced Proficient student had no better than a 2/3 chance of answering the questions correctly. Associating each page number with a Theta Value and then averaging all the responses determined the new cut-off standard for Proficient and Advanced Proficient levels.

Issues of Validity and Reliability

The traditional psychometric approach to validity considers each "type" of validity as empirical data of a given instrument. Messick (Messick, 1988) offered a Unified Approach to Construct Validity. "In the Messick approach, all of the various types of traditional psychometric validity indices are incorporated within the overall rubric of construct validity indices are based on the premise that all of the types of validity represent interrelated parts of the theoretical construct(s) underlying the measure" (Irvin et al., 2006).

High-stakes testing policies are now present in all 50 states (Tienken & Rodriguez, 2010). Little is documented and published about the validity and reliability of each state's standardized assessments. "Validity is an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions on the basis of test scores or other modes of assessment" (Messick, 1989b). Validity relates to the actual meaning that can be prescribed to a test score. It is not an inherent value or component of a test itself.

Messick (Messick, 1995) notes validity is an evolving property and part of a continuing process. The principles of validity are not exclusive to standardized assessments. They are applied to all forms of assessments from which inferences about quality and constructs should demonstrate appropriate standards of validity. Construct validity refers to the evaluative summary of both the evidence for and the actual, as well as potential, consequences of score interpretation (Messick, 1995). There are two major threats to construct validity: construct under-representation and construct-irrelevant variance.

Construct under-representation occurs when an assessment lacks the breadth needed to include all the important dimensions of a construct. Conversely, construct-irrelevant variance occurs when the assessment is too broad (Messick 1995). Construct-irrelevant variance can take one of two forms. Construct-irrelevant difficulty may occur if the reading comprehension requirements are unduly rigorous and intrude on the test's ability to measure content specific subject matter and will lead to invalidly low scores for the stated purpose of the assessment. Construct-irrelevant easiness permits individuals to respond correctly as a result of clues in the item or task, thereby creating invalidly high scores. "In its simplest terms, construct validity is the evidential basis for score interpretation" (Messick, 1995 p.743).

Messick (1995) integrated the concept of intended and unintended consequences of test score interpretation with construct validity. "Messick's proposal suggests that those who create and use high-stakes tests weigh the possible intended and unintended consequences before enacting a testing program. The integrated view of construct validity allows school administrators and policymakers to consider social and educational consequences before enacting a testing program. The integrated view of construct

validity allows schools administrators and policymakers to consider social and education consequences in the validity discussion and potentially make more informed policy decisions" (Tienken, 2011). This approach is not aligned with high-stakes testing policy where the final test score is considered the determining factor in a student's school experience.

Of equal concern is the lack of published information about specific state standardized testing programs and their conditional standard error of measurement (CSEM). Tienken (2011) notes CSEM to be an estimate of the degree of error potentially present in the test score. "Think of CSEM as the margin of error reported in political polls (e.g., + or – 7 points): The individual student-level results from every large-scale state standardized test have a margin of error…for example if a student receives a reported scale score of 546 and there are + or – 12 scale-score points of CSEM at the proficiency cut-point, then the true score could be located somewhere within the range of 534-558 and the student could be expected to score within that range if he/she took that test again" (Tienken, 2011, p. 5).

Reliability is the degree of consistency, stability, and dependability of the reported test scores (McMillan, 2004), not of the test itself. Tanner (2001) defined reliability as the "degree to which measurement data are stable" (p. 361). The authors of *The Standards of Educational and Psychological Testing* (AERA, APA, NCME, 1999) defined reliability as "the consistency of results when one repeats the testing procedure on groups or individuals." Tests with high test-scores reliability should produce similar results, if given different times.

Standards for Educational and Psychological Testing.

Messick (1995) identifies six aspects of construct validity: content-valid, substantive, structural, generalizable, external, and consequential. These aspects "function as general validity criteria or standards for all educational and psychological measurement" (Messick, 1989b).

Title 1 Law, *Improving The Academic Achievement of The Disadvantaged of the ESEA of 1965, section 10001* states: "The purpose of this law is to ensure that all children have a fair, equal, and significant opportunity to obtain a high-quality education and reach, at minimum, proficiency on challenging state academic achievement standards and state academic assessments. The purpose can be accomplished by ensuring high-quality academic assessments…In order to determine if state level standardized assessments exemplify high quality, a framework to determine quality must be clearly defined."

The Standards for Educational and Psychological Testing is produced through a long-standing collaboration of three associations: the American Educational Research Association (AERA), the American Psychological Association (APA) and the National Council on Measurement in Education (NCME). First published in 1966, *The Standards for Educational and Psychological Testing* have been revised periodically. The collaboration of the three associations has been formalized in a cooperative agreement that creates a management structure and sets procedures for maintaining and revising the Standards.

- *Technical Recommendations for Psychological Tests and Diagnostic Techniques* was prepared by a committee of the APA and published by the APA in 1955.

- *Technical Recommendations for Achievement Tests* was prepared by a committee representing AERA and the NCMUE (which is now NCME) and was published by the National Education Association in 1955.

- *Standards for Educational and Psychological Tests and Manuals* replaced the earlier two documents. It was prepared by a committee representing AERA, APA, and NCME and was published by APA in 1966.
- *Standards for Educational and Psychological Testing* was prepared by a committee representing AERA, APA, and NCME and was published by APA in 1974.
- *Standards for Educational and Psychological Testing* was a revision of the 1974 Standards. It was also a collaboration of AERA, APA, and NCME and was published by APA in 1985.
- *Standards for Educational and Psychological Testing* was a revision of the 1985 Standards. It was prepared by the Joint Committee appointed by AERA, APA, and NCME and was published by AERA in 1999. (Cronbach & Meehl 1955)

Tienken (2011) recognizes that "there exist specific statements related to construct validity, as defined by Messick (1995) and measurement error in Part I and Part II of the *Standards"* (AERA, APA, & NCME, 1999). The authors of the *Standards* concurred with Messick when they wrote: "Measurement error reduces the usefulness of measures. It limits the extent to which test results can be generalized beyond the particulars of a specific application of the measurement process. Therefore, it reduces the confidence that can be placed in any single measurement" (Tienken, 2011, p. 7). The authors also note how CSEM is of greater concern for scores near the cut-off score and

recommend CSEM be reported. By not doing so, the standardized assessment becomes less reliable and the potential for harm to students with scores near the cut-off score becomes greater.

Evidence-Based Assessment Practices

Evidence-Based Medicine (EBM) is described as the application of evidence gained from the scientific and research communities to medical practice (Guyatt & Rennie, 2002). The concept of Evidence-Based Assessment (EBA) has evolved from the practice of EBM. Similar to EBM, EBA requires empirical evidence to provide for the statistically significant use of particular assessments as treatments for particular problems in the field of behavioral sciences. While literature about EBA and its practice is most prevalent in the field of psychology, the increased presence and influence of standardized assessments in the field of education suggest the role of EBA within the field of education is worth consideration.

Impact of Demographics on Student Learning

A central piece of President Johnson's "war on poverty" policies was a commissioned study by a team of researchers led by James Coleman to survey the availability of educational opportunities for all Americans. This would become known as The Coleman Report (Coleman, Hobson, McPartland, Mood, Weinfield, & York, 1966) and remains the largest survey of public education ever undertaken. Over 640,000 students in Grades 1, 3, 6, 9, and 12, categorized into six ethnic and cultural groupsm, took achievement tests and aptitude tests, and 60,000 teachers in over 4,000 schools completed questionnaires about their background and training. Published in July of 1966, the Coleman Report concluded that schools bring little influence to bear on a child's achievement that is independent of his/her background and general social context; and that this very lack of an independent effect means that the inequalities imposed on

children by their home, neighborhood, and peer environment are carried along to become the inequalities with which they confront adult life at the end of school (Madaus, Airasian, & Kellaghan, 1980, p. 325). The Coleman Report had two primary effects on perceptions about schooling in America. First, it dealt a blow to the perception that schools could be a viable agent in equalizing the disparity in students' academic achievement due to environmental factors. Second, it spawned the perception that differences in schools have little, if any, relationship to student achievement. One of the most well-publicized findings from the Coleman Report was that schools account for only about 10 percent of the variances in student achievement; the other 90 percent was accounted for by student background characteristics (Marzano, 2000).

The findings of Coleman et al. (1966) were corroborated in 1972 when Christopher Jencks and his colleagues published *Inequality: A Reassessment of the Effects of Family and Schooling in America*, which was based on data from the Coleman report. Among the findings articulated in the study were the following:

- Schools do little to lessen the gap between rich and poor students.
- Schools do little to lessen the gap between more and less able students.
- Student achievement is primarily a function of one factor: student background
- There is little evidence that education reform can improve the influence school has on student achievement.

Coleman, et al. (1966) and Jencks (Jencks, Smith, Acland, Bane, Cohen, Gintis,, et al., 1972) demonstrate a strong correlation between student demographic data and achievement. Others have also substantiated the correlation between student demographic data and student achievement. (Alspaugh, 1991; Maylone, 2002; Payne & Biddle, 1999; Roscigno, & Ainsworth-Darnell, 1999; Sirin, 2005).

Maylone (2002) references the Standard and Poor's Statewide Evaluation Services' *Statewide Insights* study for data about the impact of SES factors on MEAP

scores. The study found a strong relationship between wealth and household income and student achievement on the MEAP. Payne and Biddle (1999) note "poor children are uniquely handicapped for education...poor homes provide little access to books, writing materials, computers, or other supports; poor students are more often distracted by diseases; they tend to live in neighborhoods affected by crime, decay, drugs and drug dealing; and their homes tend to be dysfunctional, with parents often incarcerated or disturbed" (Maylone, 2002, p. 66).

Having demonstrated that SES factors impact student achievement, Maylone (2002) went further and identified a list of district level SES factors to be paired with district MEAP scores. "If significant correlations were discovered, the factors with which they were associated were to be combined by way of multiple regression...to determine if some combination of factors (in this case, district SES) is more predictive of the independent variable (district MEAP score) than any single factor, which along with a generated constant, can be used to produce a multivariate predictive equation" (Maylone, 2002 p. 68). The sample size of the intended population of they study was 100%.

Maylone (2002) found that three district SES factors (percent of district students eligible for free- or reduced-lunch, percent of district lone-parent households, and mean annual district household income) produce a predictive equation with the most power (0.749) of a district's Composite High School MEAP Score. One limitation of Maylone (2002) is the lack of longitudinal analysis completed. The study examined only one year of MEAP scores. Maylone (2002) also demonstrated how the use of "student test scores for high stakes tend to increase the opportunity gap between students of lower SES and those of higher SES...disadvantaged students are having to look up even further to see

bottom" (Maylone, 2002, p. 103). Similar to Amrein & Berliner (2002a and 2002b) Maylone (2002) raises questions about the issue of social justice regarding high-stakes testing polices for children living in poverty.

In New Jersey, Jones (2008) aimed to create a predictive model for student achievement on the New Jersey High School Proficiency Exam (HSPA) using data published about each school district in the Annual New Jersey School Report Card. Jones (2008) went further than Maylone (2002) by analyzing the comparison between a high school's expected passing rate measured by HSPA and its actual passing rates. Jones (2008) recommends this analysis be used to determine if schools are failing to meet expectations, meeting expectations, or exceeding expectations.

Jones (2008) asked some of the following questions:

1. Which of the independent variables significantly predict the overall percentage of students in a school who will pass the Language Arts section of the New Jersey High School Proficiency Assessment (HSPA)?

2. Which of the independent variables significantly predict the percentage of students in a school who will pass the Mathematics section of the New Jersey High School Proficiency Assessment (HSPA)?

3. Are variables on the School Report Card correlated to one another?

The HSPA is one example of a high-stakes test in New Jersey because it also serves as a high school graduation requirement. This policy has been in place since the Public School Act (PSEA) of 1975 (P.L. 1975, c. 212). Each section of the HSPA includes a scoring range from 100 to 300, with a score of 199 or less equaling Partially Proficient. Scores in the range of 200 to 250 equal Proficient, and scores ranging from

250 to 300 equal Advanced Proficient. Partially Proficient scores are interpreted as failing to meet state standards for graduation. The scores for all first-time HSPA takers are included in the New Jersey School Report Card annual report. The dependent variable for Jones (2008) was the percentage of each subgroup scoring Proficient or Advanced Proficient. The independent variables included 49 data sets organized into five categories consistent with the design of the New Jersey School Report Card: school environment, student information, student performance indicators, staff information, and district financial data.

Jones (2008) found eight of the 49 variables relevant to Language Arts; HSPA scores account for nearly 90% of the variability of student achievement. The eight variables are average verbal SAT score, student mobility rate, student attendance, percentage of LEP students, percentage of students with disabilities, percentage of budget revenues from state taxes, percentage of graduates who are undecided about post-graduation plans and student attendance for Grade 11 (Jones, 2008 p. 89). A similar model was also found for math achievement on the HSPA. One limitation of Jones (2008) is the lack of control for SES or DFG data, thereby making it difficult to compare school data across the state regardless of community type. Jones (2008) does not contemplate potential validity or reliability issues with HSPA or possible unintended consequences of New Jersey and NCLB high-stakes testing policies.

Synthesis.

The aforementioned studies convincingly demonstrate the influences of district socioeconomic data on student achievement. The early studies (Coleman et al., 1966; Jencks, 1972) conclude that schools have little impact on student achievement compared

to out-of-school factors. Furthermore, Roscigno and Ainsworth-Darnell (1999) found neighborhood characteristics to be predictive of educational outcomes. This predictive power of neighborhood communities depended on the educational outcome expected. The predictions worked for student achievement scores but failed to predict high school dropout rates.

The recent studies by Maylone (2002) and Jones (2008) applied multiple regression analysis to identify specific socioeconomic variables at the district level that combine to predict student achievement. Maylone (2002) found three variables combine at the district level to predict high school MEAP scores. These variables are household income, percentage of lone-parent households, and free- and reduced-lunch eligibility. Jones (2008) required a larger mix of variables (8-9) to achieve predictive reliability for high school HSPA scores. Jones (2008) did, however, offer a helpful way to analyze high school predictive data but establishing performance expectations. This analysis could then be used to control for socioeconomic variables to identify a high school's meeting, exceeding, or failing to meet expectations. Maylone (2002) creates a need for further review of literature relevant to the specific predictive socioeconomic variables Maylone (2002) identified.

Household Income and Student Achievement.

The rare experimental studies within the field of social sciences often analyze the experience of adoptees. Plug and Vijverberg (2005) questioned whether family income matters for school outcomes by studying the achievement of adoptees and suggesting their findings from this experimental design can be interpreted as causal. Data for the study was gathered from the Wisconsin Longitudinal Survey about people born around

1939. Data collection via questionnaires began in 1957 and occurred again in 1964, 1975, and 1992. The sample size included 4,779 families with 15,726 children, including 574 adoptees.

The experimental design allowed researchers to exclude any family-related genetic influences on schooling outcomes. Plug and Vijverberg (2005) aimed to remove nature versus nurture from the debate about school outcomes. They estimated the usual relationship between family and educational outcomes for a sample of adopted children. These estimates were consistent with the idea that there is a causal relationship between family income and the educational attainment of adopted children. The study measured educational attainment based on two outcomes: the number of school years completed and college graduation

Plug and Vijverberg (2005) found that parental income has a beneficial impact on the educational attainment of adopted (genetically unrelated) children. They also found the number of siblings to have a negative impact on each child's educational attainment. Interestingly, the absence of a significant impact of parental IQ and a mother's education which are both believed to be related to parental quality, suggests that family income generates the environment in which adopted children do better in school.

Some limitations of the study are that it assumes all children are randomly given up for adoption and then randomly assigned to a new family, and it does not control for parental quality to determine its relationship to income.

Hauser and Sewell (1986) designed a study of sibling brothers, using one of the largest sample sizes found in the research, to determine family effects on education, occupational status, and earnings:

> The present analysis uses measurements of social background variables, mental ability, educational attainment, occupational status, and earnings among male Wisconsin high school graduates and a random sample of their brothers to develop and interpret simple models of socioeconomic achievement that incorporate a family variance component structure and that also correct for response variability… In the present analysis, we pool maximum likelihood estimates of models of fraternal resemblance in ability, schooling, occupational status, and earnings across two sub-samples of brother pairs from the Wisconsin Longitudinal Study (Hauser & Sewell, 1986). In one set of pairs (N = 928), we have complete data for a primary respondent but only the proxy reports about schooling and occupational status of the other brother. In the other set of pairs (N = 532), we have complete, self-reported data for both members of each pair, plus the proxy reports about the other brother (Hauser & Sewell, 1986, p. 85).

With a significant sample size and over 90 percent of surveys returned, the size of Hauser and Sewell's study (1986) is noteworthy. The study found family background to have large independent effects on ability, schooling, and occupational attainment. This was a common theme in the literature about the influence of household income on student achievement. It matters, but not as much as other family background characteristics.

A large body of quantitative research concerned with the impact of household income on student achievement applies economic production theories to estimate effect sizes of specific variables.

Blau (1999) studied the effect of parental income on a child's cognitive, social, and emotional development. To measure these three areas, Blau (1999) analyzed data

generated from six different assessments:

- Peabody Individual Achievement Tests (PIAT)
- Reading Recognition
- Peabody Picture Vocabulary Test
- Verbal Memory Parts, A and B
- Behavior Problems Index
- Motor and Social Development

Blau (1999) found the effect of current income on child development is small while the effect of permanent income is larger but not as influential as family background characteristics. The outcome with the largest income effects were the Behavior Problems Index score suggesting child behaviors are more malleable and influenced by household income than aptitude. Blau (1999) concludes the effect size of permanent household income is not large enough to suggest income transfers policies to be feasible or beneficial. To significantly impact student achievement through household income it would require an unprecedented transfer of income. Rather, Blau (1999) notes family background and other family and child characteristics have a greater effect on child development than household income.

Weinberg (2001) built upon earlier research (Blau, 1999), which established a positive relationship between parental income and child educational attainment by analyzing the relationship of income to parental behaviors. Weinberg (Weinberg, 2001) corroborates that parental income is related to student outcomes but also demonstrates that parental income impacts child-rearing practices. The greater the parental income,

the greater a parent's ability to mold a child's behavior through pecuniary incentives, while lower parental incomes increase reliance on corporal punishment practices in the home.

Synthesis

Existing literature differs about the degree to which household income influences student achievement. Coleman et al. (1966) and Jencks (1972) conclude that many out-of-school factors influence student achievement, with household income as one of the factors identified. More recent research relies on more sophisticated statistical analysis to isolate the power of the influence of specific out-of-school variables on student achievement. There is agreement in the literature across disciplines that the influence of family background on student outcomes is greater than family income. There is also agreement that the impact of family income is greater in lower-income households than in higher-income households. Of note, literature relating to lone-parent households and student achievement further highlights the influence of household income on mediating the potential negative effects of the lone-parent or step-household.

Lone-Parent Household and Student Achievement

By 1990, it had been well established that students from a single-parent household were less likely to complete high school or attend college (Amato & Keith, 1991; Coleman, 1988). A closer examination of both the lone-parent household and step-household impact on student achievement became warranted.

Peterson and Zill (Peterson & Zill, 1986) examined data from a 1981 national sample of 1400 children between the ages of 12 and 16 to determine the effects of marital disruption on child behavior. While negative effects of divorce are associated with

negative child outcomes (increased levels of depression, stress, anxiety, aggression and emotional disturbance), these negative effects were found to be lower if the child lives with the same-sex parent or at least maintains a good relationship with the same-sex parent.

Dawson (1991) represents a foundational study about the impact of lone-parent households on student achievement by examining the impact of divorce and step-households on several student behaviors. Dawson (1991) studied data from the 1998 National Health Interview Survey on Child Health. The sample size for the study included 17,110 children under the age of 18 living with single biological mothers or biological mothers and a stepfather. Dawson (1991) found these children were more likely to repeat a grade of school, be expelled, be treated for emotional or behavior problems, and be vulnerable to health problems than children living with both biological parents. Dawson's (1991) findings conclude the lone-parent household with biological mother has similar effects as the step-household on student achievement and overall well-being.

Amato and Keith (Amato & Keith, 1991) conducted a meta-analysis of the influence of parental divorce and its association with the education attainment of children. This study calculated effect sizes for 15 outcome variables across 37 studies involving 81,000 individuals. Mean effect sizes were significant and negative for all outcomes. Educational attainment was defined by two factors: high school graduation and overall number of years of schooling. Amato and Keith (1991) found that adults who experience divorce experience lower levels of overall well-being than those who do not.

Similar to Dawson's (1991) findings, parental divorce is associated with lower educational attainment.

Following the work of Dawson (1991) and Amato and Keith (1991), studies began to isolate specific characteristics of step-households and their effects on student achievement. Astone and McLanahan (1991) examined the relationship between specific family structures and high school achievement. By distinguishing between children in single-parent homes and those living in step-households, they allowed for the effect of number of parents to be isolated to determine whether the number of adults in a household is a critical factor in student success in high school. The sample of the study included 1000 randomly selected U.S. high school students surveyed in 1982, 1984, and 1986. The study examined three questions:

1. Are children who live with single parents and stepparents exposed to different parental expectations and styles?

2. Are school-related parenting practices related to student achievement?

3. Can differences in parent behavior account for any of the negative associations between family structure and student achievement? (Astone & McLanahan, 1991)

Children from non-intact families (lone-parent or step-household) report lower educational expectations on the part of their parents, less monitoring of schoolwork, and less overall supervision of social activities. Children from non-intact families were more likely to be disengaged from school. Stepparent families had similar negative consequences as single-parent families.

The bulk of research demonstrates that children in step-households perform less well than those from mother/father households and often closely resemble

the achievement levels of youths from single-parent households...children in step-households had lower school class ranks and were more likely to repeat a grade than children in households containing both mother and father. Although their family incomes were higher, children in step-households were similar in performance to children in single-parent households (Downey, 1995).

Downey (1995) further examined children in step-households by studying a national sample of eighth grade students from the 1988 National Education Longitudinal Study. The sample size included 24,599 eighth graders, within which 1,192 were living in mother/stepfather households and 470 were living in father/stepmother households. Education attainment measured included report card grades and standardized test scores for math, reading, science, and history. Also of note about the design of the study, data were collected from two teachers and one parent about each student. The study asked two questions:

1. Do parental resources mediate the effect on student achievement of living in a step-household?

2. Does the sex of the biological parent or child affect this process?

Downey (1995) concluded that much of the difference in academic achievement for students in step-households can be explained by parental cultural, economic, and interpersonal resources. The results also suggest student achievement does not appear to be significantly affected by the sex of the biological parent living with the child. Downey (1995) supports earlier findings about the negative impact on student achievement for students in lone-parent households and step-households. Also of note, income was found to effectively mediate aspects of these negative effects.

More recently, Carlson and Corcoran (Carlson & Corcoran, 2001) examined data from the National Longitudinal Survey of Youth to determine the effect of various family structures on behavioral and cognitive outcomes for children ages 7 to 10. Regarding behavioral outcomes. it was found that family income, mother's psychological functioning, and the quality of home environment combine to significantly impact behavior. Regarding cognitive outcomes, family income and mother's aptitude were found to significantly impact the child's cognitive development. Behavior problems were measured with the Behavior Problem Index (BPI) created by Zill and Peterson (1986). Cognitive ability was measured with the Peabody Individual Achievement Test (PIAT) for math and reading.

> With respect to our theoretical framework, we find that family structure does operate through economic status because, once income is controlled, the family structure effects primarily appear for both behavioral and cognitive outcomes. Average family income is much lower for the non-intact family configurations than for two-parent families and is lowest for children raised in continuous single-parent families. Family income itself strongly predicts most cognitive and behavioral outcomes.....Overall, this study highlights the importance of examining multiple categories of family structure, of analyzing effects across more than one outcome domain, and of evaluating a range of theoretical mechanisms that may mediate between family structure and child well being. (Carlson & Corcoran, 2001, p. 791).

Synthesis

The literature concerning the influence of lone-parent household on student achievement suggests that examination of these effects should be disaggregated between lone-parent household and step-households. Interestingly, the literature in this area provided further evidence to support the significant influence family income plays on mediating the potential negative effects of either household situation. With the exception of Peterson and Zill (1986) the literature did not assign significant positive influence to lone-parent households when both parent and child share the same gender. Indeed, much of the literature found the gender relationship, when isolated, to have little to no significant effect on student achievement or well-being. The literature suggests a distinction between lone-parent households and step-households is not necessary. Instead, a clear delineation between a married-couple household and lone-parent household is sufficient. This will impact the theoretical framework of this study.

Poverty and Student Achievement

Sirin (2005) conducted a meta-analysis of 74 independent studies published between 1990 and 2000 with the aim of determining the relationship between socio-economic factors (SES) and academic achievement. The sample size included 101,157 students from 6,871 schools in 128 school districts. Sirin (2005) aimed to replicate White (1982), and consistent with previous studies, Sirin (2005) found a medium to strong relationship between socioeconomic variables and student achievement at the school level. The relationship became weaker at the individual student level. Sirin (2005) also identified flaws with White (K. R. White, 1982), including that (a) White (1982) allowed for multiple correlations from the same sample, violating the principle that there can be

only one unique correlation from one unique samples, (b) the sample was not limited to U.S. schools, and (c) White accepted IQ scores as a measure of student achievement.

Sirin (2005) found the magnitude and degree of the relationship between SES and academic achievement to be contingent upon several factors. Unlike White (1982) and Coleman et al. (1966), Sirin (2005) concluded the relationship between SES and academic achievement increases across levels of school from primary through middle school with the exception of high school. For example, neighborhood SES factors were more predictive of student achievement than family SES factors. Sirin (2005) notes a decrease in the overall strength between SES factors and student achievement compared with the findings of White (1982). Sirin (2005) highlights the importance of the unit of analysis when measuring the influence of SES factors on variables by noting the following: following:

1. SES is a multi-dimensional construct in which different components yield different results.
2. Free/reduced lunch data is problematic when used to determine SES factors.
3. SES has more meaning for minority students.
4. SES is limited in its capacity to capture student social and economic background. (Sirin, 2005)

> Of all the factors examined in the meta-analytic literature, family SES at the student level is one of the strongest correlates of academic performance. At the school level, the correlations were even stronger. This review's overall finding, therefore, suggests that parents' location in the socioeconomic structure has a strong impact on students' academic achievement. Family SES sets the stage for

students' academic performance both by directly providing resources at home and by indirectly providing the social capital that is necessary to succeed in school (Coleman, 1988, p. 438).

The relationship of SES and student achievement has been the subject of much study. The largest studies (Coleman et al. 1966; Jencks 1972) have received the most attention and scrutiny. Both were discussed earlier in the paper. As noted, Sirin (2005) suggests that school and community resources, or lack thereof, influence student achievement. Pereira (2011) notes, "In 1972 Mosteller and Moynihan stated that it was their belief that one of the most significant findings of the Coleman Report (1966) was that there was very little difference between the resources allocated to Black and White students, therefore claiming the gaps in achievement are the direct results of some other factor....Jencks et al. also found significance in other results brought forth by Coleman and his colleagues, such as the academic achievement improvement of students with lower socioeconomic backgrounds that attend schools with affluent peers (Pereira, 2011, p. 56).

The question therefore arises: Do school resources matter and can they mitigate the negative effects of lower household income, lone-parent household and parent education levels? Jencks (1972) suggests a family background's strong influence on student achievement, as well as other SES factors, renders school resources virtually powerless. Instead, Jencks (1972) concludes that the great determining factor of student success to be "luck."

Hanushek (Hanushek, 1997; Hanushek, 1986) found no relationship between school resources and student achievement. Rather, Hanushek (1986) concludes that

schools operate with high levels of inefficiency and therefore the allocation of additional resources would be unwise and wasteful.

Greenwald et al. criticizes Hanushek (1986), explaining, "Hanushek's synthesis method, vote counting, consists of categorizing, by significance and direction, the relation between school resource inputs and student outcomes (included but not limited to achievement)" (p. 362) (Pereira, 2011, p. 57). Pereira (2011) notes, "Greenwald et al. criticized Hanushek's vote counting method, identifying it as an outdated, rather insensitive, procedure for summarizing results....After conducting a reanalysis of Hanushek's (1986) conclusions, Greenwald et al. affirmed the data on the relation between school resources inputs and student outcomes, including achievement, were substantially more consistent and positive than he believed" (p.362) (Pereira, 2011 p.57).

Greenwald (Greenwald, Hedges, & Laine, 1996a) and Hanushek (Hanushek, 1986) analyze similar data and arrive at conflicting conclusions. Pereira (2011) provides a thorough analysis of both Greenwald's (Greenwald et al., 1996a; Greenwald, Hedges, & Laine, 1996b) and Hanushek's (Hanushek, 1996) responses and cites Gamoran and Long (Gamoran & Long, 2006) as an objective third party which determined, "these conflicting reports are the result of the researchers' difference in inclusion criteria when selecting studies for their analyses; Greenwald et al. was more selective, whereas Hanushek classified findings of previous studies as negative, positive, or neutral. Greenwald et al. and Hanushek may have differed in some aspects of their findings; however, these researchers did agree that (a) in at least some cases, higher levels of resources are associated with higher achievement, (b) the qualities of schools that produce these effects are hard to pin down, and (c) the ways in which resources are used

is more consequential for achievement than the presence or absence of resources" (Gamoran & Long, 2006, p. 8) (Pereira, 2011, p. 63).

Synthesis

For the purpose of this study, the findings of Sirin (2005) and Gamoran and Long (2006) suggest school resources alone fail to consistently account for student achievement. Indeed, out-of-school variables must be controlled for when determining a school district's influence on student achievement. Furthermore, the disagreement between Hanushek (1986, 1996) and Greenwald et al. (1996a & 1996b), summarized brilliantly by Pereira (2011), demonstrated how school resources alone did not consistently account for significant influence on student achievement. Sirin (2005) concluded that family SES at the student level accounts for the greatest influence on student achievement. Previously reviewed literature of this study identified household income and lone-parent household status to be significant indicators of family SES. The literature also referenced the role of parental education on student achievement. This aspect warrants further consideration.

Parental Education and Student Achievement

Previously cited work has noted the influence of parental education levels on student achievement. The literature on achievement consistently has shown that parent education is important in predicting children's achievement (Davis-Kean, 2005; Jimerson, Egeland, & Teo, 1999; Klebanov, Brooks-Gunn, & Duncan, 1994; Kohn, 1963; Luster, Rhodes, & Hass, 1989; Smith, Brooks-Gunn, & Klebanov, 1997)

Kandel and Lesser (1969) investigated the relationship between the influence of parents on a student's education goals versus the influence of peers under different social

and cultural conditions. During the spring of 1965, data were collected for all students from three high schools (N=2377) from surveys. Next, data was collected on the mothers from 60% of the population. The schools were representative of a large, lower class urban community, a rural community and a diverse regional high school. The empirical data "provide strong evidence for Kahl's (1953) observation that parental aspiration is a more important determinant of children's educational aspirations than is social-class membership per se. This does not deny the importance of social class as a determining factor in educational aspirations, but our data can be interpreted to show that the impact of social class on the adolescent, to the extent that it exists, is absorbed in the nature of the maternal influence" (Kandel & Lesser, 1969, p. 218). Also of consequence, the study found that class played a significant role in both the educational goals and encouragement mothers provide children regarding educational attainment. Middle-class mothers were found to provide more encouragement and have higher expectations for educational attainment than mothers raising a family with lower-class means. "These parental attitudes and plans, in turn, are associated with social-class position. But for the child, the parent is clearly the link between social-class and position and future life goals" (Kandel & Lesser, 1969, p. 220).

More recently, Davis-Kean (2005) examined a broad national cross-sectional sample of 868 8 to 12 year olds (436 females, 433 males; 49% non-Hispanic European American, 47% African American). Davis-Kean (2005) examined the power of the indirect role of parental expectations on the home environment and found the following: "The results suggest that the amount of schooling that parents receive influences how

they structure their home environment as well as how they interact with their children in promoting academic achievement" (Davis-Kean, 2005, p. 300).

Synthesis

There is little disagreement within the literature about the influence of parental education levels on student achievement. Davis-Kean (2005) represents the most recent thinking about these influences by identifying a link between parental education levels and parenting beliefs and behaviors. These findings suggest parental level of education to be a significant aspect of SES family factors influencing student achievement.

Importance of Early Childhood Learning and Intervention

Maylone (2002) and Jones (2008) found predictive formulas for high school level student achievement data through regression analysis of various district demographic data. A growing body of research demonstrates the importance of early childhood learning and how achievement gaps between subgroups are present by kindergarten and first grade. Jencks and Phillips found one half of the Black/White achievement gap in reading and mathematics in Grade 12 could be explained by differences in first grade scores (Jencks & Phillips, 1998). The Early Childhood Longitudinal Study Kindergarten Cohort of 1998 conducted by the Department of Education found a gap between the skills White and Black students have upon entering school (Fryer & Levitt, 2004). This gap exists and continues to grow each year, even when controlling for socioeconomic factors. Furthermore, the academic skills of economically disadvantaged students are significantly lower than their more affluent peers (Fryer & Levitt, 2004).

Nagaoka & Roderick (2005) found the achievement gap for students begins as early as first grade and recommends that schools need to invest in effective early

assessment programs for reading intervention. A landmark meta-analysis (Lipsey & Wilson, 1993) of more than 300 studies from various disciplines found significant evidence of the "general efficacy" of early intervention childhood programs. Shonkoff and Meisels (2000) state, "During the years from 3 to 10, the brain is more densely wired than at any other time in a child's life…for language acquisition and the process to facilitate logical thinking" (Meisels & Shonkoff, 2000, p. xi). Carnegie (1996) demonstrated how remediation of skills becomes more difficult over time.

More recently, in a study funded by the Annie E. Casey Foundation, Hernandez (2011) found:

- One in six children who are not reading proficiently in third grade do not graduate from high school on time, a rate four times greater than that for proficient readers.
- The rates are highest for the low, below-basic readers; 23% of these children drop out or fail to finish high school on time, compared to 9% of children with basic reading skills and 4% of proficient readers.
- Overall, 22% of children who have lived in poverty do not graduate from high school, compared to 6% of those who have never been poor. This rises to 32% for students spending more than half of their childhood in poverty.
- For children who were poor for at least a year and were not reading proficiently in third grade, the proportion that don't finish school rose to 26%. That is more than six times the rate for all proficient readers.
- The rate was highest for poor Black and Hispanic students, at 31 and 33% respectively, or about eight times the rate for all proficient readers. Even

among poor children who were proficient readers in third grade, 11% still did not finish high school. That compares to 9% of subpar third grade readers who have never been poor.

- Among children who have never lived in poverty, all but 2% of the best third grade readers graduated from high school on time.
- Graduation rates for Black and Hispanic students who were not proficient readers in the third grade lagged far behind those for White students with the same reading skills (Hernandez, 2011).

Research clearly suggests early intervention is more likely to positively impact student achievement. While the predictive models demonstrated by Maylone (2002) and Jones (2008) raise questions about the equitable impact of high-stakes testing policies, they do little to provide educators with data to impact instructional decisions when opportunity still exists to lessen achievement gaps.

Theoretical Framework

Schools are not able to control out-of-school influences on student achievement. A large research base demonstrates that out-of-school factors significantly impact student learning (Alspaugh, 1991; Amato & Keith, 1991; Astone & McLanahan, 1991; Blau, 1999; Coleman, Hobson, McPartland, Mood, Weinfield, & York, 1966; Davis-Kean, 2005; Dawson, 1991; Downey, 1995; Hauser & Sewell, 1986; Jencks, Smith, Acland, Bane, Cohen, Gintis, et al., 1972; Payne & Biddle, 1999; Peterson & Zill, 1986; Plug & Vijverberg, 2005; Roscigno & Ainsworth-Darnell, 1999; Sirin, 2005; Weinberg, 2001).

Existing literature about the use of high-stakes standardized assessment data to determine the quality and success of a school district can be categorized by two

contradictory conclusions. One group of studies concluded high-stakes accountability systems prove to be an effective reform policy and have improved student achievement over the past two decades (Braun, 2004; Hanushek & Raymond, 2003)These studies also concluded that these policies need to be improved in several areas. Another group of studies concluded high-stakes accountability systems have done little to improve student achievement and in some cases have done harm (Amrein & Berliner, March 2002; Amrein & Berliner, December 2002; Elmore, 2004; Linn, 2000; Mehrens, 1998; Rustique-Forrester, 2005). There is general agreement within the research that high-stakes accountability systems have not impacted the achievement gap between economically disadvantaged students and their more affluent peers (Hanushek, Raymond, & Rivkin, 2004). There is also evidence to suggest this achievement gap was significantly reduced during the 1970s and early 1980s before increasing again during the 1990s (Lee, 2002).

Also of note, state level high-stakes standardized assessments have limitations and flaws (Tienken, 2011). Therefore, it cannot be assumed high-stakes data generated from standardized assessments accurately measure the quality and success of a school district. Consequently, education policymakers may be operating under the false assumption that high scores on high-stakes standardized assessments accurately identify quality and success in school districts. Education policymakers may be rewarding or punishing school districts based on a false paradigm by using high-stakes test data to identify quality and success in school districts without consideration or control for other significant socioeconomic variables proven to impact student achievement as measured by standardized assessments.

Furthermore, a large body of research demonstrates the importance of effective early childhood education and early interventions ((Fryer & Levitt, 2004; Jencks & Phillips, 1998; Lipsey & Wilson, 1993; Nagaoka & Roderick, 2005). Therefore, an imperative exists to more accurately identify New Jersey school districts exceeding expectations at the elementary level while controlling for out-of-school socioeconomic variables. Empirical data is needed to determine the predictive validity of school district socioeconomic data on student achievement as measured by high-stakes standardized assessments.

Recent studies have shown how multiple regression analysis of district level socioeconomic data and student achievement data can be used to determine a predictive formula of district level student achievement (Jones, 2008; Maylone, 2002). In New Jersey, the present system for categorizing schools districts based on socioeconomic data does not include predictive data which could be used to better determine which districts are meeting, exceeding, or failing to meet expectations while controlling for each district's socioeconomic data.

The Essentialist paradigm of public policy relevant to public schooling has a weak research base and fails to account for or acknowledge influences of outside factors on student achievement as measured by standardized tests scores. The Progressive paradigm has a strong research base and accepts that a comprehensive view of student achievement is required with standardized test data considered as one of many factors.

The line of inquiry for this study aimed to establish a research base for a new policy context for evaluating school district effectiveness by more accurately determining

which districts are effectively impacting student achievement when controlling for out-of school district variables most predictive of student achievement.

The problem this study seeks to examine is the lack of control for district socio-economic and economic demographics when evaluating school district standardized Grade 3 NJ ASK assessment data for student achievement in Language Arts and Mathematics. Furthermore, no study has examined the impact of these demographics on standardized assessment data at the elementary school level in New Jersey. Therefore, this study examines student achievement in Grade 3, the first grade level tested by the NJ ASK, in Language Arts and Mathematics. The Literature Review examined the following: The Essentialist Paradigm, the Progressive Paradigm, Review of Assessment and High-Stakes Policy Development, Technical Characteristics of Standardized Assessments, Importance of Early Childhood Learning and Intervention, and Influence of Demographic Factors on Achievement.

Sirin (2005) showed family SES is the most important determinant of school financing because in the United States half of all public school funding is based on the property taxes within each school district. State and federal subsidies fail to create equitable funding across school districts and communities. Based on current school financing policies, a situation is created where students from lower SES families are most likely to attend school districts that are at best financially inferior to wealthier school districts.

Maylone (2002) found predictive power of student achievement by combining percentage of lone-parent household, mean annual district household income, and percentage of free- and reduced-lunch at the high school level. The literature suggests

(Haveman, R. & Wolfe, 1995; Sirin, 2005) free- and reduced-lunch information can be problematic as an identifier for the family effects of SES. Sirin (2005) notes the ES (effect size) for SES influence on student achievement increases with each grade level during the primary and middle school years and then the ES decreases in high school. This finding provides further evidence for the design of this study to focus on third grade scores because this is where the SES influence will be in its earliest stage.

The extant literature reviewed suggests the variables of household income, percentage of lone-parent households, and level of parental education within a school district may combine to explain and predict student achievement as measured by standardized tests.

Production Function Theory

For the purpose of this study, the district is considered the institution. The inputs become the district variables identified previously. The output is the district NJ ASK 3 scores in Language Arts and Mathematics. This study design is an empirical research design with its roots in microeconomics. Hanushek (1986) notes, "The economics research on schooling is empirical in nature and an understanding of its findings must begin with an underlying conceptual model of the educational process. A natural starting point is the economic model of production theory and firm behavior" (Hanushek, 1986). By definition, a function associates one quantity with another. When applying the production function theory to social science, it is implied one or more inputs (independent variables) influence the output (dependent variable).

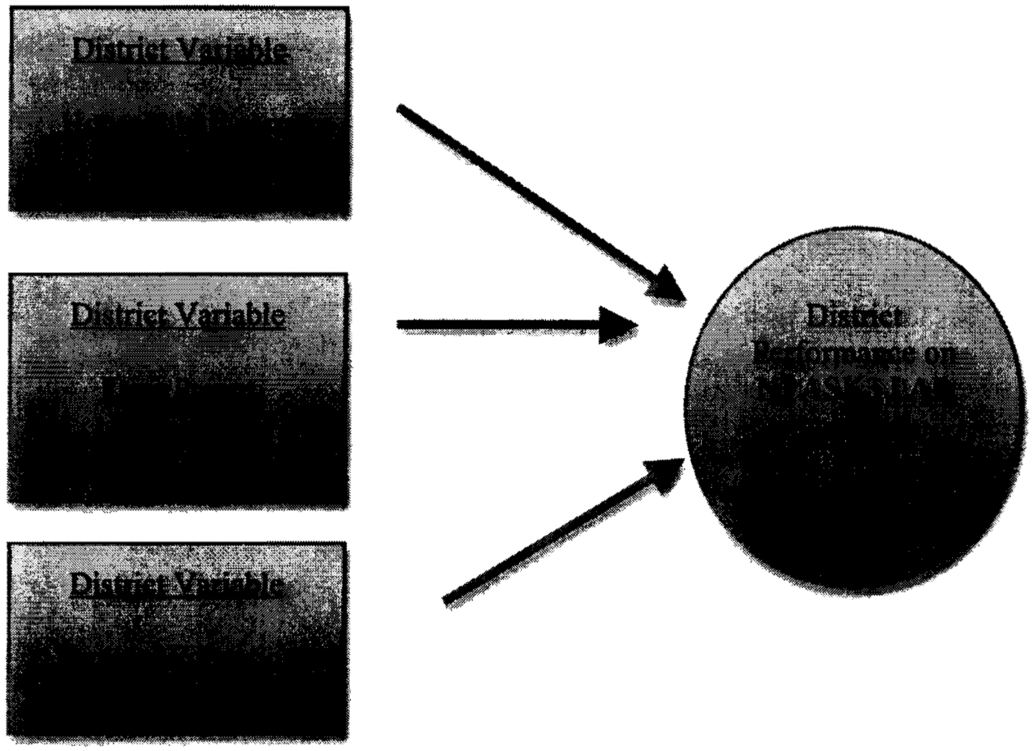


Figure 1. Input/Output Framework

Campbell's Law

The Essentialist paradigm continues to dominate American educational policy, as evidenced by the proliferation of high-stakes testing policies in all fifty states. These policies continue to place increasing importance on quantifiable measures of student achievement and, more recently, teacher effectiveness. Consequently, a principle of social science known as Campbell's Law is worthy of consideration within the field of education. Campbell's Law states:

> The more any quantitative social indicator is used for social decision making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor. (Berliner & Nichols, March 2005)

Nichols and Berliner (2005) apply this principle to conclude "the over-reliance on high-stakes testing has serious negative repercussions that are present at every level of the public school system" (Berliner & Nichols, 2005, p.1). They identified ten categories of evidence within public education where Campbell's Law could be applied due to the presence of high-stakes testing. The ten categories are:

- Administrator and Teacher Cheating
- Student Cheating
- Exclusion of Low-Performing Students From Testing
- Misrepresentation of Student Dropouts
- Teaching to the Test
- Narrowing the Curriculum
- Conflicting Accountability Ratings
- Questions about the Meaning of Proficiency
- Declining Teacher Morale
- Score Reporting Errors

Chapter Summary

Thomas Jefferson stated the purpose of American public education was to "develop an intelligent citizenry and to provide educational opportunities that guarantee each individual the chance for optimal development" (Tanner & Tanner, 2007, p. 4). During the dominance of the Progressive Era, *The Cardinal Principles of Secondary Education* offered another answer. The report concluded, "Education in a democracy, both within and without the school, should develop in each individual the knowledge, interests, ideals, habits, and powers whereby he will find his place and use that place to

shape both himself and society toward ever nobler ends" (National Education Association of the United States, Commission on the Reorganization of Secondary Education, 1918, p. 3). This report laid the groundwork for a massive expansion of the public school system by legitimizing both the expectation that every student should graduate from high school and the value of educating the "whole" child. Both Jefferson and *The Cardinal Principles of Secondary Education* articulated a vision of a public school consistent with the tenets of Progressivism. Today the Common Core Standards state the purpose of public education is to prepare every student for success in college, a much more limited view of the public school system than Jefferson or *The Cardinal Principles* recommended and further evidence for the dominance of Essentialist theory. Interestingly, the Eight-Year Study and Thorndike (1924) provide a significant research base for how Progressive strategies have proven to effectively prepare students to achieve the Essentialist vision of the public school. Present day high-stakes standardized assessment policies, which are rooted in Essentialist theory, lack a significant research base to suggest they will be effectively prepare students for success in college.

One group of studies about high-stakes accountability systems concluded that they prove to be an effective reform policy and have improved student achievement over the past two decades (Braun, 2004; Hanushek & Raymond, 2003; Hanushek & Raymond, 2004; Hanushek & Raymond, 2005). These studies also concluded that these policies need to be improved in several areas. Another group of studies concluded that high-stakes accountability systems have done little to improve student achievement and in some cases have done harm (Amrein & Berliner, March 2002; Amrein & Berliner, December 2002; Dorn, 1998; Elmore, 2004; Linn, 2000; Mehrens, 1998; Rustique-

Forrester, 2005). There is general agreement within the research that high-stakes accountability systems have not impacted the achievement gap between economically disadvantaged students and their more affluent peers (Hanushek, Raymond, & Rivkin, 2004). There is also evidence to suggest this achievement gap was significantly reduced during the 1970s and early 1980s before increasing again during the 1990s (Lee, 2002).

The Literature Review outlined the evolution of two dueling paradigms of thought about American public education over the past century. Theorists were given particular attention for the extent of their influence on later theorists and the volume of their writing. These theorists/researchers include William Bagley, Edward Thorndike, Frederick Taylor, Francis W. Parker, and John Dewey. Next, a review of two landmark studies about the influences of out-of-school variables on student achievement provided insight about how district socio-demographic data influences district student achievement. A more thorough examination of the literature revealed how the variables of household income, parental education, family poverty and lone-parent households significantly influence student achievement. Amrein and Berliner (2002a and 2002b) provide the context for examining the debate about the influence of high-stakes testing policy on student achievement on NAEP scores at the state level.

The most ardent and prolific defender of high-stakes testing policies, Raymond and Hanushek (2003), attempted to discredit Amrein and Berliner (2002a and 2002b). Pereira (2011), however, artfully notes, "Although Raymond and Hanushek attempted to disprove Amrein and Berliner's (2002a) theories of negative high-stakes influences on student learning, what they did instead was emphasize just how insignificant the point advantages truly were. Acknowledging that the NAEP mathematics scores range on a 0

to 500 scale, even at the greatest recorded point advantage of 5.3 for the Grade 4 test during the 1992-2000 period, that is till only a 1.06% increase. Therefore, this "advantage" proves completely insignificant to use as evidence of a positive correlation between increased student achievement and the implementation of high-stakes attachment to tests, as Raymond and Hanushek maintained that "rigorous analysis reveals that accountability policies have had a positive impact on test scores during the past decade" (2003, p. 50) (Pereira, 2011, p. 49). Pereira (2011) demonstrated the thin body of research to support the overall effect of these polices.

Considering the lack of research to support the effectiveness of Essentialist theories, skepticism about Essentialist policies is warranted. Indeed, the most telling data analysis of NAEP scores may be by Mintrop and Sunderman (Mintrop & Sunderman, 2009). The achievement gap trend lines of NAEP scores follow more of a curvilinear pattern, with the closing of achievement gaps in the 1970s and early 1980s followed by an increase in the late 1980s and 1990s with no significant progress since the enactment of NCLB in 2002. A watershed moment for Essentialist theories is marked by the publication of *A Nation at Risk* (National Commission on Excellence in Education, 1983). Can it simply be coincidental the increased influence of Essential thinking over education policy coincided with the reversal of gains made in closing the achievement gap as evidenced by NAEP scores?

Maylone (2002) found the predictive power of student achievement by combining percent of lone-parent household, mean annual district household income, and percentage of free- and reduced-lunch at the high school level. The literature suggested (Haveman, R. & Wolfe, 1995; Sirin, 2005) free- and reduced-lunch information can be problematic

as an identifier for the family effects of SES. Sirin (2005) notes the ES (effect size) for SES influence on student achievement increases with each grade level during the primary and middle school years and then the ES decreases in high school. This finding provided further evidence for the design of this study to focus on third grade scores because this is where the SES influence will be in its earliest stage.

The extant literature reviewed suggested the variables of household income, percentage of lone-parent households, and the level of parental education within a school district may combine to explain and predict student achievement as measured by standardized tests.

Chapter III

METHODOLOGY

The purpose for this study was to identify which specific school community demographic factors account for the greatest amount of variance in a New Jersey school district's percentage of students scoring Proficient or above on the 2009 NJ ASK 3 in Languages Arts and Mathematics. The study intentionally limited its focus to out-of-school variables on district NJ ASK 3 data. If out-of-school variables are found to explain significant variance in district test scores or even predict a district's scores, as the existing literature suggests, the value of using district test scores to measure the quality of in-school variables may be in question.

Research Design

This study used a non-experimental correlational, explanatory, cross-sectional design with quantitative methods. Within the field of social sciences, most research problems are not easily examined through experimentation. Therefore, correlational studies are one of the frequently used research designs in the social sciences and can limit research from finding causality between two variables. Non-experimental causal-comparative research designs do attempt to provide evidence of cause and effect relationships between variables and can be seen as a non-experimental research design which may identify causality. Johnson (2002) suggests "there is no reason to believe a stronger causal claim can be made from a study with controls measuring the relationship between gender and test grades (a causal-comparative study) than from a study without controls measuring the relationship between time spent studying for a test and test grades (a correlational study)." Johnson (2002) contends causal-comparative research is neither

better nor worse than correlational research in determining causality between two variables. Johnson (2002) proposes a new classification of nonexperimental research into three categories: descriptive research, predictive research, and explanatory research. This approach classifies research based on the stated objective of a research study. "It is also helpful to classify nonexperimental quantitative research according to the time dimension…here the types of research include cross-sectional research, longitudinal research, and retrospective research" (Johnson, 2001). Because this study aims to explore the relationship between two or more variables from one moment in time with quantitative methods, it is appropriately designed as a non-experimental, correlational, explanatory, cross-sectional study:

This study builds upon aspects of Maylone (2002) and Jones (2008).

> A prediction study is an attempt to determine which of a number of variables are most highly related to the criterion variable. Prediction studies are conducted to facilitate decision making about individuals, to aid in various types of selection, and to determine the predictive validity of measuring instruments. Typical prediction studies include those used to predict an individual's likely level of success in a specific course (e.g., first-year algebra), those that predict which students are likely to succeed in college or in a vocational training program, and those that predict the area of study in which an individual is most likely to be successful. Thus, the results of prediction studies are used not only by researchers but also by counselors, admissions directors, and employers….More than one variable can be used to make predictions. If several predictor variables correlate well with a criterion, then a prediction based on a combination of those variables

will be more accurate than a prediction based on any one of them (Hanushek, 1986, p.203).

Multiple linear regression models were used to determine the statistical significance of out-of-school variables on school district 2009 NJ ASK Grade 3 Language Arts and Mathematics scores. The community variables presented in Chapter II were identified in the literature as influencing student achievement measured by standardized assessments and provided the basis for the theoretical framework of the study. The strength of these variables' relationship to school district 2009 NJ ASK Grade 3 scores was unknown.

Hinkle, Wiersman, & Jurs (2003) suggest, "The behavioral and social sciences could not exist without statistics. Behavioral scientists use statistics to explain the results of research studies and to provide empirical evidence to support or refute theories" (Hinkle, Wiersman, & Jurs, 2003, p. 2). When examining data, if the characteristics of data are the same, then these characteristics are considered constant. When characteristics of data are different, then these characteristics are considered variables. This study examined three different independent variables and their influence and predictive power on one dependent variable through multiple linear regressions.

"In multiple linear regression, we have a single criterion variable (Y) and multiple predictor variables (Xi,). The multiple regression equation contains a regression coefficient (bi) for each predictor variable and the regression constant (a) "(Hinkle et al., 2003, p. 462).

This study examined the following independent variables:

1. Household Income, defined as:
   - Median district household income
   - Percentage of families below poverty
   - Percentage of economically disadvantaged
   - Percentage of household annual income under $30,000
   - Percentage of household annual income above $200,000
4. Lone-Parent Household, defined as combination of:
   - Percentage of district male households, no wife
   - Percentage of district female households, no male
5. Parental Education, defined as:
   - Percentage of 25 years or older, no high school diploma
   - Percentage of 25 years or older, high school graduate
   - Percentage of 25 years or older, high school graduate, some college experience
   - Percentage of 25 years or older, bachelor's degree
   - Percentage of 25 years or older, advanced degree

The dependent variables for this study were school district 2009 NJ ASK 3 Language Arts and Mathematics scores, which were defined as the percentage of the student population that achieved either a Proficient or Advanced Proficient score.

Data for each school district's median household income was taken from the American Community Survey portion of the 2010 U.S. Census. American FactFinder was used to locate the data. The survey generated three different sets of data: one-year

estimates, three-year estimates and five-year estimates. This study examined five-year estimates because they provided the largest sample size.

The first variable identified by the literature to influence student achievement as measured by standardized assessments was the percentage of mean district household income. Furthermore, this study aimed to replicate aspects of Maylone (2002) where mean district household income was found to combine with the percentage of district lone-parent households and the percentage of district students eligible for free- and reduced-price lunches to predict MEAP scores. This study examined the following data relevant to household income:

1. Median District Household Income
2. Percentage of Families Below Poverty
3. Percentage of Economically Disadvantaged
4. Percentage of Household Income under $30,000
5. Percentage of Household Income above $200,000

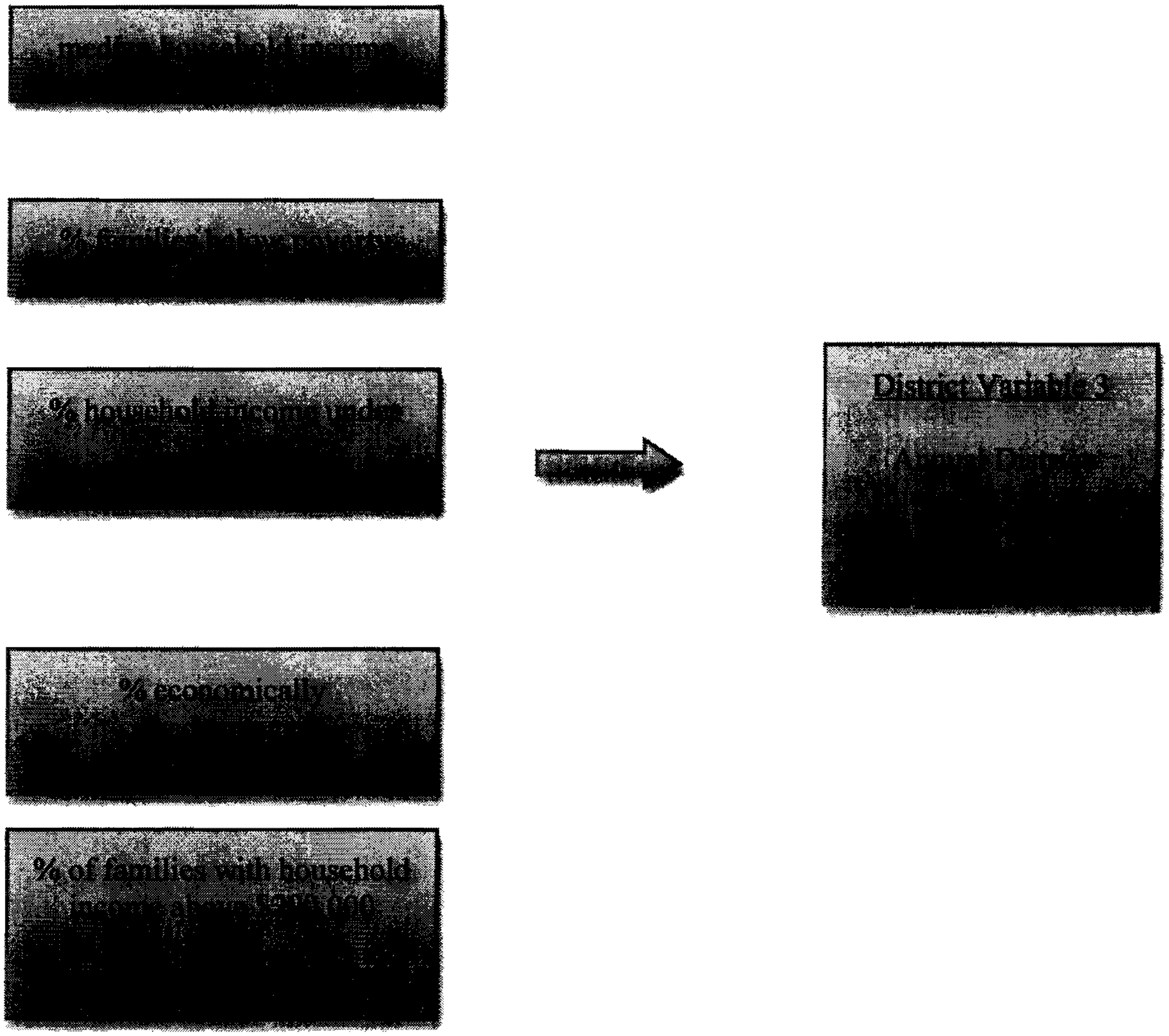


Figure 2. Annual District Household Income Construct

The second variable identified by the literature to influence student achievement as measured by standardized assessments was the percentage of lone-parent households in school districts. Maylone (2002) also found this variable to have significant predictive power of student achievement as measured by standardized assessments. Data about this district variable was also taken from the American Community Survey. The literature demonstrated there was no significant difference in influence between lone-parent

households and households with a stepparent. Therefore, for the purpose of this study, the following data points from the American Community Survey were combined to determine district lone-parent household:

1. Male householder, no wife present
2. Female householder, no husband present

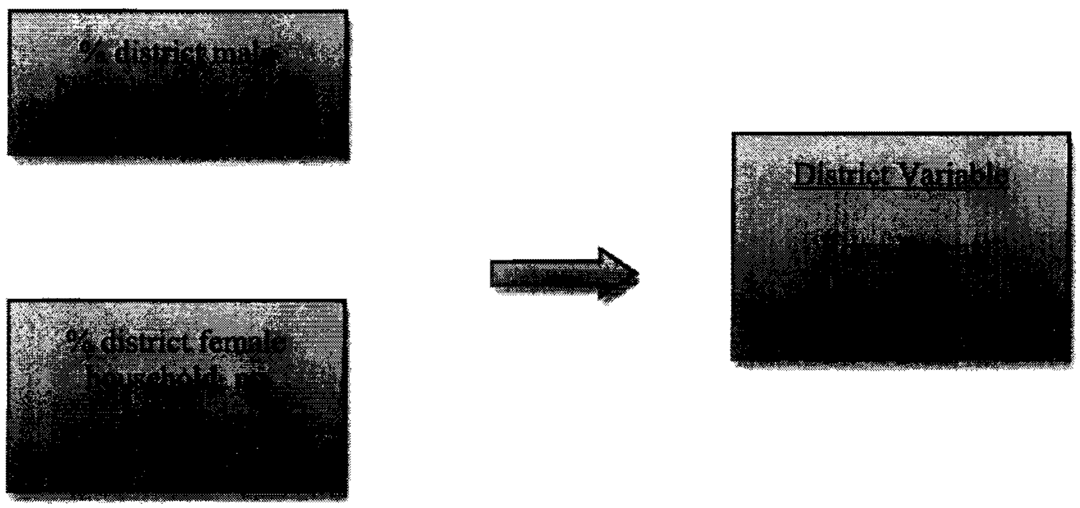


Figure 3. Annual District Percentage Lone Parent Household Construct

The third variable identified by the literature to influence student achievement as measured by standardized assessments was percentage level of parental education in school districts. Maylone (2002) found the third predictive variable to be percent of district students eligible for free- and reduced-price lunches to significantly predict student achievement as measured by standardized assessments. As reviewed in Chapter II, Sirin (2005) discovered significant issues with relying on free- and reduced-price lunches to determine influences of socioeconomic factors on student achievement. The

literature did suggest that the level of parental education significantly influenced student achievement. Therefore, for the purpose of this study, the level of parental education was examined as another independent variable.

Data about the level of parental education was taken from the American Community Survey. This data was organized into five categories:

1. Level of Parental Education – Percentage 25 years or older, no high school diploma
2. Level of Parental Education – Percentage 25 years or older, high school graduate
3. Level of Parental Education – Percentage 25 years or older, high school graduate with some college experience
4. Level of Parental Education – Percentage 25 years or older, bachelor's degree
5. Level of Parental Education – Percentage 25 years or older, advanced degree

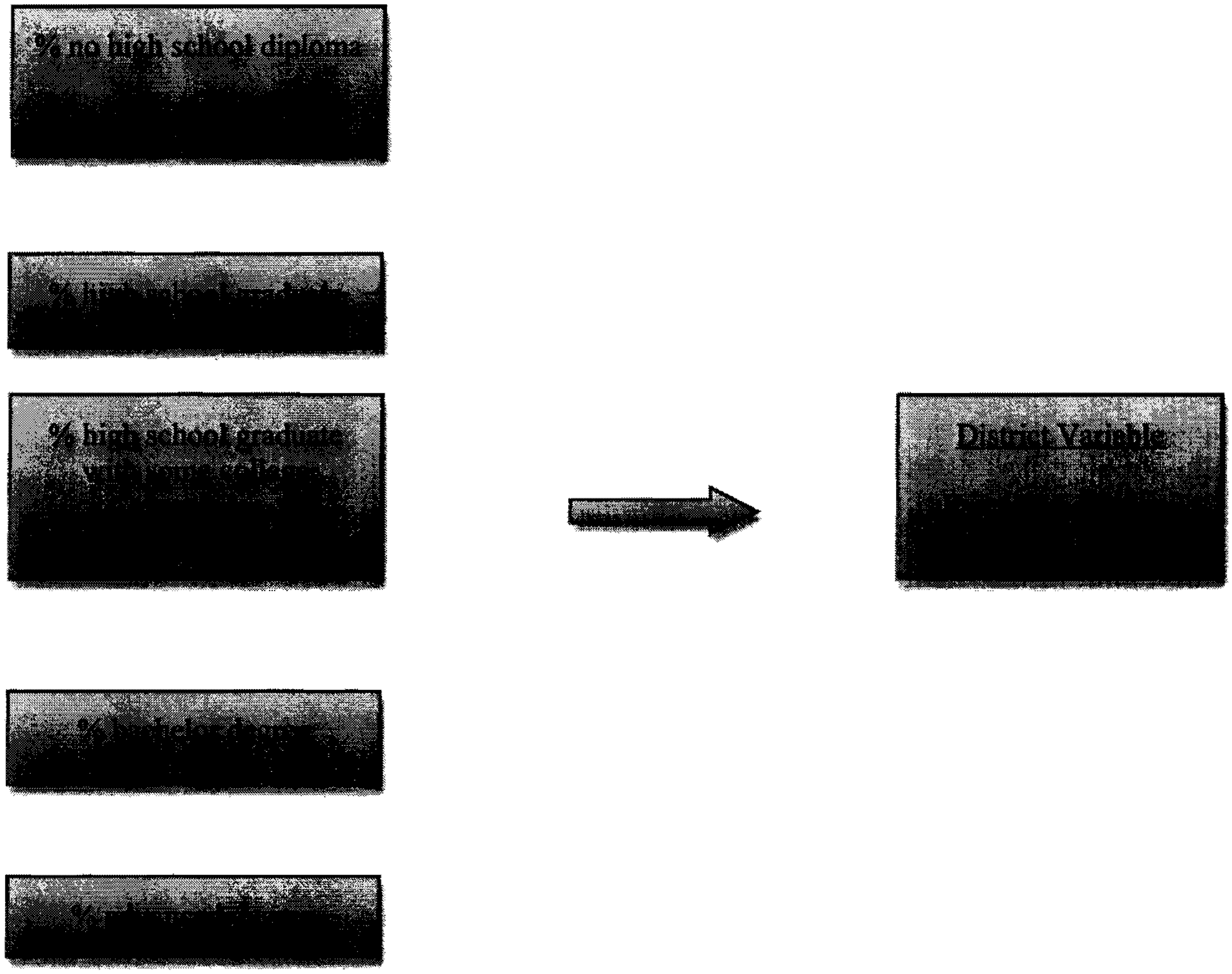


Figure 4. Annual District Percentage Parental Education Construct

Research Questions

This study examined three overarching research questions:

1. How much variance in NJ ASK 3 2009 test results in Language Arts and Mathematics is explained by out-of-school socioeconomic variables?

2. How accurately can out-of-school socioeconomic and community-level variables predict a school district's percentage of students scoring Proficient or above on

the NJ ASK 3 Language Arts and Mathematics sections?

3. Which combination of community-level variables account for the greatest amount of variance in a school district's percentage of students passing the NJ ASK 3?

To gain a deeper understanding about these questions, and after a thorough review of extant literature, eight research questions were developed:

Research Question 1: How much variance in 2009 NJ ASK 3 test results in Language Arts can be explained by the household income construct (Table 2) for New Jersey school districts?

Research Question 2: How much variance in 2009 NJ ASK 3 test results in Mathematics can be explained by the household income construct (Table 2) for New Jersey school districts?

Research Question 3: How much variance in 2009 NJ ASK 3 test results in Language Arts can be explained by the lone-parent household construct (Table 3) for New Jersey school districts?

Research Question 4: How much variance in 2009 NJ ASK 3 test results in Mathematics can be explained by the lone-parent household construct (Table 3) for New Jersey school districts?

Research Question 5: How much variance in 2009 NJ ASK 3 test results in Language Arts can be explained by the percentage of parental education construct (Table 4) for New Jersey school districts?

Research Question 6: How much variance in 2009 NJ ASK 3 test results in Mathematics can be explained by percentage of parental education construct (Table 4) for New Jersey school districts?

Research Question 7: Which combination of independent variables establishes the greatest reliable predictive power for a school district's 2009 NJ ASK 3 Language Arts test results?

Research Question 8: Which combination of independent variables establishes the greatest reliable predictive power for a school district's 2009 NJ ASK 3 Mathematics test results?

Null Hypotheses

Null Hypothesis 1: No statistically significant relationship exists between 2009 NJ ASK 3 scores in Language Arts and the household income construct for New Jersey school districts.

Null Hypothesis 2: No statistically significant relationship exists between 2009 NJ ASK 3 scores in Mathematics and the household income construct for New Jersey school districts.

Null Hypothesis 3: No statistically significant relationship exists between 2009 NJ ASK 3 test results in Language Arts and the percentage of lone-parent household construct for New Jersey school districts.

Null Hypothesis 4: No statistically significant relationship exists between 2009 NJ ASK 3 test results in Mathematics and the percentage of lone-parent household construct for New Jersey school districts.

Null Hypothesis 5: No statistically significant relationship exits between 2009 NJ ASK 3 test results in Language Arts and the percentage level of parental education construct for New Jersey school districts.

Null Hypothesis 6: No statistically significant relationship exits between 2009 NJ ASK 3 test results in Mathematics and the percentage level of parental education construct for New Jersey school districts.

Null Hypothesis 7: There is no statistically research demonstrated combination of independent variables with reliable predictive power for 2009 NJ ASK 3 test results in Language Arts for New Jersey school districts.

Null Hypothesis 8: There is no statistically research demonstrated combination of independent variables with reliable predictive power for 2009 NJ ASK 3 test results in Mathematics for New Jersey school districts.

Population

All school district data examined in this study related to Grade 3 student achievement as measured by 2009 NJ ASK 3 for Language Arts and Mathematics. Presently, New Jersey has approximately 572 school districts categorized into eight different district factor group (DFG) categories determined by 2010 U.S. Census data. The categories listed from districts located in the state's poorest communities to districts located in the state's wealthiest districts are as follows: A, B, CD, DE, FG, GH, I, J. Some of the 572 school districts in New Jersey include regional high schools and county schools that do not serve elementary grades. The target population for this study was 100% of all New Jersey school districts with both 2009 NJ ASK 3 data and 2009 Census data with at least 25 students enrolled in Grade 3. Therefore the available population for the study was 438 districts and the sample size for the study was 438 school districts: 100% of the population.

Data Collection

Data about the dependent variables of 2009 NJ ASK 3 Language Arts and Mathematics scores for New Jersey school districts was readily available through the annual publication of the New Jersey School Report Card. Three scores of proficiency percentages are reported for both Language Arts and Mathematics results: Partial Proficient, Proficient, and Advanced Proficient. For the purpose of this study, Proficient and Advanced scores were combined to indicate one passing rate. The data was downloaded directly from the New Jersey Department of Education website into an Excel spreadsheet, where it could be more easily analyzed alongside the data for the independent variables. (http://www.nj.gov/education/schools/achievement/2010/njask3/).

The data for the independent variables in this study were gathered from two locations. Data about the percentage of economically disadvantaged families in each district was also downloaded into an Excel spreadsheet from the New Jersey Department of Education website, where the annual School Report Card data is stored (http://education.state.nj.us/rc/2009/index.html).

Data about the remaining independent variables for each New Jersey school district were gathered from the American Community Survey (ACS). This nationwide survey is one aspect of the U. S. Census Bureau's decennial census program.

> The ACS began in 1996 and has expanded each subsequent year. Full nationwide implementation began in January 2005 for housing units and in January 2006 for group quarters (GQ). Starting with the 2005 ACS, one-year estimates have been available for geographic areas with populations of 65,000 or more. In 2008, the first ACS three-year estimates were released for geographic areas with

populations of 20,000 or more. For small areas (less than 20,000 population), it will take five years to accumulate a large enough sample to provide estimates with accuracy similar to the decennial census. Beginning in 2010, and every year thereafter, the nation will have these five-year period estimates available, a resource that shows change over time, even for neighborhoods and rural areas (American FactFinder, December 2010, p. 1).

The 2005-2009 ACS 5-year estimates were used for this study and based on data collected between January 2005 and December 2009. While less current than the one-year and three-year estimates the five-year estimates represent a larger sample size and are published for small geographic areas, which include areas with population under 20,000. The data was downloaded into Excel spreadsheets from the following website: http://factfinder.census.gov/servlet/DatasetMainPageServlet?_program=ACS&_submenuId=datasets_1&_lang=en&_ts=

Instrumentation

Instrumentation for this study included district level scores on 2009 NJ ASK 3 in Language Arts and Mathematics. The study aimed to determine the amount of variance in 2009 NJ ASK 3 scores explained by out-of-school district variables.

Reliability

"Reliability is the degree to which a test consistently measures what it is measuring. The more reliable a test is, the more confidence we can have that the scores obtained from the test are essentially the same scores that would be obtained if the test were re-administered to the same test takers at another time or by a different person. If a

test is unreliable...then scores will likely be quite different every time the test is administered" (L. R. Gay, Mills, & Airasian, 2009, p. 158).

NJ ASK Grade 3-8 Technical Report of 2009 stated the following:

> In reading this technical report, it is critical to remember that the testing program does not exist in a vacuum; it is not just a test. It is one part of a complex network intended to help schools focus their energies on dramatic improvement in student learning. NJ ASK is an integrated program of testing, accountability, and curricular and instructional support. It can only be evaluated properly within this full context (New Jersey Department of Education, 2009b, p. 12).

Although the NJDOE recommends district leaders view NJ ASK scores within a "full context" when making decisions about a student's educational experience, Tienken (2008) found New Jersey education leaders across the state do not hed this warning. In fact, school leaders, among other decisions, are using NJ ASK data to "stream students into basic skills instruction and Title I programs (elementary. middle school) and recommend remedial high school course sequences, partially or totally depending on the district, on state results" (Tienken, 2008, p. 56).

Pereira (2011) notes the theoretical foundation for NJ ASK assessments was Classical Test Theory (CTT). "The foundation of CTT is built upon the ideals that a total test score is comprised of multiple items. The CTT approach assumes 'that the raw score (X) obtained by any one individual is made up of a true component (T) and a random error (E) component: X=T+E (Kline, 2005, p.91). Taking a person's mean scores on the same test providing they had an infinite number of testing sessions would be the only manner in which one may obtain a person's true score. Since that is an impossibility, the

central aspect of CTT is T, although this number is merely hypothetical" (Pereira, 2011, p. 142).

NJ ASK scores are used to make high-stakes decisions about school districts, schools, teachers, and students. Therefore, decision makers should know the standard error measurement for each assessment. Tienken and Rodriguez (2010) explain one reason for the significant variance of standard error of measurement (SEM) in individual test scores to be the number of questions on a test used to measure student understanding of a particulate standard. There are simply too few.

Since the dependent variable of this study was NJ ASK 3 Language Arts and Mathematics scores for New Jersey school districts, the reliability of the results must be considered within the context of standard error of measurement (SEM) as reported by NJDOE.

In determining the degree of reliability for specific score results, a reliability coefficient is calculated to determine how reliable a measure is (Reinard, 2006). Reinard (2006) notes:

Reliability coefficients should be as close to 1.00 as possible. However, interpretations often are based on guidelines such as the following:

.90 and above -- highly reliable

.80 -.89 -- good reliability

.70 -.79 -- fair reliability

.60 - 69 -- marginal reliability

under 60 -- unacceptable reliability

Pereira (2011) demonstrated the importance of analyzing content sections clusters of each portion of the NJ ASK test. Table 1 displays the coefficient alpha and SEM for NJ ASK 3 2009 Language Arts and Mathematics clusters.

Table 1

Coefficient Alpha of NJ ASK 3 and Standard Error Measurement for Multiple Choice Clusters and Short-Constructed Responses

Subject/Cluster	No. of items	Alphas	SEM
LAL Writing	2	.70	1.42
LAL MC	18	.82	1.74
Working with Text	12	.78	1.36
Analyzing Text	6	.54	1.07
Math SCR*	6	.61	1.02
Math MC	35	.88	2.46
Number & Numerical Operations	14	.80	1.49
Geometry & Measurement	7	.45	1.12
Patterns & Algebra	7	.60	1.10
Data Analysis, Probability & Discrete Mathematics	7	.66	1.09
Problem Solving	18	.82	1.74

*Math SCR refers to the open-ended math response question graded with a rubric.

Based on the coefficient reliability scale provided by Reinard (2006), the clusters scores for Analyzing Text and Geometry & Measurement were unreliable. The cluster scores for Patterns &Algebra, Data Analysis, Probability & Discrete Mathematics, and Math SCR were marginally reliable. However, this study used full test scale scores, and full test reliability is between the acceptable range of .80 -.89 for looking at large groups. These results will be considered greater detail in Chapter IV of this study.

Validity

"Validity refers to the degree to which a test measures what it is supposed to measure and, consequently, permits appropriate interpretation of scores. Validity is, therefore, 'the most fundamental consideration in developing and evaluating tests (American Psychological Association, 1999 p. 9) When we test, we test for a purpose, and our measurement tools must help us achieve that purpose" (Gay et al., 2009, p. 151).

When evaluating assessment validity, research generally looks at four different measures: content validity, criterion-related validity, construct validity, consequential validity.

Table 2. Forms of Validity

Form	Method	Purpose
Content Validity	Compare content of the test to the domain being measured	To what extent does this test represent the general domain of interest?
Criterion-Related Validity	Correlate scores from one instrument of scores on a criterion measure, either at the same (concurrent) or different (predictive) time.	To what extent does this test correlate highly with another test?
Construct Validity	Amass convergent, divergent, and content-related evidence to determine that the presumed construct is what is being measured	To what extent does this test reflect the construct it is intended to measure?
Consequential Validity	Observe and determine whether the test has adverse consequence for test takers of users.	To what extent does the test create harmful consequences for the test taker?

(Gay et al., 2009 p.151)

Regarding content validity and the appropriateness of the defined content, the *NJ ASK 2009 Technical Report* states:

The review process required by the State Board involved teachers, school

administrators, students, parents, and representatives from business, higher education, and the community. In addition, several content areas were reviewed by Achieve, Inc., and the Council of Chief State School Officers (CCSSO). In response to this unprecedented review, the 2004 New Jersey Core Curriculum Content Standards provide the level of specificity and depth of content that will better prepare students for post secondary education and employment. The standards are based on the latest research in each of the content areas and identify the essential core of learning for all students. Since the adoption of the original 1996 New Jersey Core Curriculum Content Standards (CCCS), the New Jersey State Board of Education approved administrative code that implements all aspects of standards-based reform. N.J.A.C. 6A:8 requires districts to align all curriculum to the standards; ensure that teachers provide instruction according to the standards, ensure student performance is assessed in each content area, and provide teachers with opportunities for professional development that focuses on the standards. (New Jersey Department of Education, 2009a p. 186)

To ensure that the NJ ASK 2010 test design meets standards of adequacy for content representation, the following process was followed:

Adequate representation of the content domains defined in the common core content standards (CCCS) is assured through use of a test blueprint and a responsible test construction process. New Jersey performance standards, as well as the CCCS, are taken into consideration in the writing of multiple-choice and constructed response items and constructed-response rubric development. Each test must align with and proportionally represent the sub-domains of the test

blueprint....The CCCS are represented on each test by balancing sub-domain coverage on each test, by proportionally representing items corresponding to Partially Proficient, Proficient, and Advanced Proficient performance categories on each test, and by matching item format to the requirements of the content and standards descriptions (New Jersey Department of Education, 2009a p.186).

The construct validity of NJ ASK 3 2009 was monitored by "studying patterns of relationships to provide evidence supporting the inferences made from test scores....The correlations between clusters within a content area were generally found to be higher than the correlations between clusters across the content areas (New Jersey Department of Education, 2009a, p.187).

To determine the extent to which the content of NJ ASK 3 2009 correlates with previous NJ ASK 3 test administrations, the following statistical analysis was completed:

> For each administration, classical item analyses are completed prior to item calibration, scaling, and equating. These statistics are calculated again once all of the data are available. These analyses involve computing a set of statistics based on classical test theory for every item in each form. Each statistic is designed to provide some key information about the quality of each item from an empirical perspective. The statistics estimated for the NJ ASK 3 are described below.
>
> Classical item difficulty ("P-Value"): This statistic indicates the percentage of examinees in the sample that answered the item correctly. Desired p-values generally fall within the range of 0.30 to 0.90.
>
> Item discrimination ("r-biserial"): This statistic is measured by the poly-serial correlation between the item score and the test criterion score and describes

the relationship between performance on the specific item and performance on the entire form. Higher values indicate greater differences in the performance of competent and less competent examinees. Items with negative correlations can indicate serious problems with the item content (e.g., multiple correct answers or unusually complex content), or can indicate that students have not been taught the content. For LAL, the test criterion score is the total score of all reading items (MC and CR) and the writing prompt. For mathematics, the test criterion score is the total score of all MC and CR (Extended Constructed Response (ECR) and Short Constructed Response (SCR) items. For science, the test criterion score is also the total score of all MC and CR items.

The proportion of students choosing each response option: These statistics indicate the percentage of examinees that select each of the available answer options and the percentage of examinees that omitted the item. (d) Distractor analyses for MC items: This statistic reports the percentage of examinees who select each incorrect response (distractor).

Percentage of students omitting an item: This statistic is useful for identifying problems with test features such as testing time and item/test layout. Typically, we would expect that if students have an adequate amount of testing time, 95% of students should attempt to answer each question. When a pattern of omit percentages exceeds 5% for a series of items at the end of a timed section, this may indicate that there was insufficient time for students to complete all items. Alternatively, if the omit percentage is greater than 5% for a single item, this could be an indication of an item/test layout problem. For example, students

might accidentally skip an item that follows a lengthy stem (New Jersey Department of Education, 2009a, p. 65).

The Technical Report for the NJ ASK 2009 Grades 3-8 Assessment does not address the question of consequential validity. Chapter II of this study indicates this may be an area worthy of consideration. Tienken (2009) notes, "Educators should not use the NJ ASK 3…scores as the only factor for making decisions about student entrance into specific programs (e.g. Basic Skills, Title I, Gifted and Talented, Honors courses). The technical characteristics for the test results and the inherent social justice issues cannot justify the possible negative consequences attached to their use in a high-stakes manner. The confluence of sub-domain reliability estimates, relationships between district factor group and student test results, and sizeable standard error of measurement creates a conundrum for educators" (Tienken, 2009, p. 58).

Methods

Step One: Collection of Data

The review of literature identified the following independent variables to establish the household income construct, lone-parent household construct and parental education construct:

- Percentage of economically disadvantaged students
- Percentage of population 25 years or older without a high school diploma
- Percentage of population 25 years or older with a high school diploma
- Percentage of population 25 years or older with a high school diploma and some college
- Percentage of population 25 years or older with a bachelor's degree

- Percentage of population 25 years or older with an advanced degree
- Percentage of lone-parent households
- Percentage of households with less than $30,000 income in the past 12 months
- Percentage of households with more than $200,000 income in the past 12 months
- Percentage of families below poverty level
- Median family income for the district

The percentage of economically disadvantaged students for each district was taken from the NJDOE school report card database, which was available at the NJDOE website. NJDOE defines economically disadvantaged students as students eligible for free- or reduced- price breakfast or lunch meals.

The level of parental education construct for each district required the combination of the following categories reported out by ACS:

- less than 9th grade = no high school diploma
- 9th to 12th grade = no high school diploma
- high school graduate = high school diploma
- some college, no degree = high school diploma, with some college
- associates degree = high school diploma, with some college
- bachelor's degree = college graduate
- graduate or professional degree = advanced degree graduate

To determine the percentage of lone-parent households for each district the following categories of ACS data were combined:

- Male householder, no wife present, family
- Female householder, no husband present, family

The dependent variables of percentage passing for NJ ASK 3 Language Arts and Mathematics were calculated by combining the percentage of Proficient and Advanced Proficient scores for each district. The dependent variables for this study were:

- Percentage of students passing 2009 NJ ASK 3 Language Arts
- Percentage of students passing 2009 NJ ASK 3 Mathematics

All data used for this study were archival, public information, and located via the Internet on sites such as the New Jersey Department of Education and Data Universe and Zip Skinny. No permissions were required to access the data. The data for independent variables and the two dependent variables were entered into one Excel spreadsheet. Each row in the spreadsheet represented a different New Jersey school district. The five-year estimates of each independent variable were reported as total numbers, not percentages. Once all the data were combined into the single Excel spreadsheets, additional columns were added to calculate the total percentages of students passing for each district by dividing the district number by the total population for each category and multiplying by one hundred.

Step Two: Alignment of Data

Because data were pulled from two different databases, attention was paid to data alignment before the databases were combined. This alignment was done in two phases. First, the demographic data and district 2009 NJ ASK 3 data were opened simultaneously in two different Excel spreadsheets. The data were sorted first by district code and then by district name. This allowed for a side-by-side comparison to verify which districts had demographic data and 2009 NJ ASK 3 data. Only districts that met the criteria of having both types of data were included in the study. After the alignment was complete,

481 of 590 total school districts in New Jersey were found to meet the criteria for inclusion. Next, school districts with less than 25 students enrolled in Grade 3 were eliminated (n=25), which brought the total population to 438.

The second phase of alignment required the merging of both databases. Before the merge, 20 data sets were selected at random from the demographic database to audit for accuracy. For example, Robbinsville Township was selected and the row number and district code were verified to be the same in each set of data. This process was repeated for each data set to verify alignment with the district 2009 NJ ASK3. After the 20 data sets were verified; the data were merged into one database and the process was repeated to check for accuracy using the same 20 districts.

After all the data for one dependent variable were merged into one spreadsheet, additional columns were created in which all the independent variable district demographic data were converted into a percentage for every school district. At this point all columns except the columns with the percentage data for the independent variables were hidden from view to aid in readability when working with the data set.

Step Three: SPSS Data Entry, Examination and Outputs

After the data were merged into one Excel spreadsheet, the spreadsheet was uploaded into SPSS (Statistical Package for the Social Sciences), where a correlational analysis was done for each dependent variable. Correlation coefficients were generated for each independent variable to determine the level of strength and direction of the relationship between each independent variable and the dependent variable. Scatter plots were created to determine if any irregularities might disqualify a dependent variable as an indicator of the relationship with the independent variable.

Step Four: Multiple Regressions

After examination of the correlational outputs, three different linear multiple regression models for each dependent variable were completed. The first relied on the stepwise method, the second applied the theoretical framework established from the review of extant literature, and the third tested the theoretical framework through hierarchical linear regression. All consider the threat of multicollinearity on the predictive variables in constructing their models. The structure of this analysis is explained in Chapter IV.

The stepwise multiple regressions for the dependent variable of 2009 NJ ASK 3 LAL scores produced five models. The linear multiple regressions for dependent variableS of 2009 NJ ASK 3 Math scores produced six models. The theoretical framework simultaneous regression model produced one model for each dependent variable. Stepwise, simultaneous, and hierarchical regression analyses were run to better understand the threat of multicollinearity on the predictive power of each model.

Step Five: Application of Predictive Formula

Maylone (2002) found the numerical coefficients for three SES factors that predicted 56% of the variance in his sample:

1. Percent of students eligible for free- or reduced-price lunches
2. Percent of district lone-parent households
3. Mean annual district household income

These three values were combined with the SPSS generated numerical constant to create the following predictor equation:

$$-0.226a + -0.767b + 0.00014c + 64.533 =$$

a = Percent of students eligible for free- or reduced-price lunches

b = Percent of district lone-parent households

c = Mean annual district household income

For the purpose of this study, a similar algorithm was created for each of the 13 models identified through the stepwise regression method and the theoretical framework. This algorithm was then applied to 100% of the population in a new column labeled Predictive Model with the model number. Another column was added next to the Predictive Model labeled Difference (Diff.) between predicted and actual scores. The actual 2009 NJ ASK 3 score for each district was then subtracted from the predicted score. The result was entered as the Diff. score. Last, the standard deviation of the differences was calculated for all 13 models and entered at the bottom of each Diff. column.

Analysis Strategy

The data for this study was analyzed by creating one database for each dependent variable. The total population for the study included 438 school districts. Simultaneous multiple linear regression, stepwise regression, and hierarchical linear models were then created by importing each dependent variable database into the IBM SPSS (Statistical Package for the Social Sciences) predictive analytics software. A two-way ANOVA (analysis of variance) was generated for each dependent variable. The F-static was analyzed to determine if each regression model was statistically significant. To determine which model explained the greatest variance in each dependent variable,.0 an analysis of each model's Adjusted R^2 (coefficient of determination) was conducted. The main purpose of Adjusted R^2 is to determine how well variables in the model will predict

future outcomes based on their explanation of variance in the dependent variable. Within each model the independent variables reported a standardized beta coefficient, which was used to compare the strength of the effect of each independent variable on the dependent variable within each statistically significant model.

One of the most significant threats to the reliability and validity of the linear regression models was the impact of multicollinearity on the independent variables. Multicollinearity occurs when more than one of the predictor variables in a multiple linear regression model is highly related. For example, the percentage of families earning less than $30,000 annual income and the percentage of students classified as economically disadvantaged in a single school district is likely to be highly related. While multicollinearity doesn't impact the overall predictive power of a regression model, it can cause individual coefficient estimates to change erratically. This can negatively impact calculations regarding the predictive power of individual school districts. Since a major aspect of this study included the application of the formula created by Maylone (2002) to individual school districts, multicollinearity had to be given serious consideration.

Tolerance and Variance Inflation Factor (VIF) values for each predictor in the model are used to measure the degree of multicollinearity between predictive variables in a multiple regression model. Tolerance is the reciprocal of VIF. Tolerance = $1 - R^2$ where VIF = 1/ tolerance. A VIF less than 5 is considered a high standard to verify that multicollinearity does not significantly impact predictor variables in a multiple regression model.

To verify that multicollinearity did not threaten the predictive reliability of the multiple regression models generated for each dependent variable, three different methods were used to create the regression models. First, stepwise regression was applied to each dependent variable and all of the independent variables. This method sequences variables based on F-tests and Tolerance levels to build models with the greatest R^2 values and lowest multicollinearity levels. Second, the theoretical framework established through review of extant literature was applied to build simultaneous multiple regression models for each dependent variable. The condition for the application of the theoretical framework to predictive variables was the following:

- One variable from each construct must be used (household income, lone-parent household, level of parental education)
- VIF for all three variables must be less than 2.

The model which produced the highest R^2 while meeting the above two conditions was identified as the best theoretical framework model.

Third, the three independent variables identified as the best model from the theoretical framework simultaneous multiple regressions were run through hierarchical linear regression. The independent variable with the highest beta was entered first, followed by the independent variable with the second highest beta, and then the independent variable with the lowest beta was entered. Hierarchical linear regression was conducted on these three variables to thoroughly examine the potential impact of multicollinearity on the model's predictive power.

Last, The betas and constant from the stepwise method and theoretical framework regressions were applied to the formula created by Maylone (2002) and added to the

database for each dependent variable. These formulas generated a predicted score for each school district. The actual score was then subtracted from the predicted score and the total labeled as the difference. The difference was calculated for each dependent variable for each school district. Finally, the standard deviation was calculated for the distribution of the differences for each predictive formula. The predictive formula which generated the smallest standard deviation in the differences was considered to be the best predictive model for each dependent variable.

Chapter Summary

This study used a non-experimental, correlational, explanatory, cross-sectional design with quantitative methods. Within the field of social sciences, most research problems are not easily examined through experimentation. Therefore, correlational studies are one of the frequently used research designs in the social sciences and can limit research from finding causality between two variables. Non-experimental causal-comparative research designs do attempt to provide evidence of cause and effect relationships between variables and can be seen as a non-experimental research design which may identify causality. Johnson (2002) suggests "there is no reason to believe a stronger causal claim can be made from a study with controls measuring the relationship between gender and test grades (a causal-comparative study) than from a study without controls measuring the relationship between time spent studying for a test and test grades (a correlational study)." Johnson (2002) contends causal-comparative research is neither better nor worse than correlational research in determining causality between two variables. Johnson (2002) proposes a new classification of nonexperimental research into three categories: descriptive research, predictive research, and explanatory research.

This approach classifies research based on the stated objective of a research study. "It is also helpful to classify nonexperimental quantitative research according to the time dimension...here the types of research include cross-sectional research, longitudinal research, and retrospective research" (Johnson, 2001). Because this study aims to explore the relationship between two or more variables from one moment in time with quantitative methods, it is appropriately designed as a non-experimental, correlational, explanatory, cross-sectional study.

This study examined the following independent variables:

1. Household Income, defined as:
    - Median district household income
    - Percentage of families below poverty
    - Percentage of economically disadvantaged students
    - Percentage of household annual income under $30,000
    - Percentage of household annual income above $200,000
2. Lone-Parent Household, defined as a combination of:
    - Percentage of district male households, no wife
    - Percentage of district female households, no male
3. Parental Education, defined as:
    - Percentage 25 years or older, no high school diploma
    - Percentage of 25 years or older, high school graduate
    - Percentage of 25 years or older, high school graduates, some college experience
    - Percentage of 25 years or older, bachelor's degree

- Percentage of 25 years or older, advanced degree

The dependent variables for this study were school district 2009 NJ ASK 3 Language Arts and Mathematics scores, which were defined as the percentage of the student population that achieved either a Proficient or Advance Proficient score. The total sample size for the study was 438 school districts, which comprised 100% of the population.

Multiple linear regression models were used to determine the statistical significance of out-of-school variables on school district 2009 NJ ASK Grade 3 Language Arts and Mathematics scores. After examination of the correlational outputs, three different linear multiple regression models for each dependent variable were completed. The first relied on the stepwise method, the second applied the theoretical framework established from the review of extant literature, and the third tested the theoretical framework through hierarchical linear regression. All consider the threat of multicollinearity on the predictive variables in constructing their models.

The community variables presented in Chapter II were identified in the literature as influencing student achievement measured by standardized assessments and provided the basis for the theoretical framework for this study. The strength of the variables relationship to school district 2009 NJ ASK Grade 3 scores in Language Arts and Mathematics was unknown.

Chapter IV

FINDINGS

The purpose for this study was to identify specific school community demographic factors that account for the greatest amount of variance in a New Jersey

school district's percentage of students scoring Proficient or above on NJ ASK 3 in Languages Arts and Mathematics. The study intentionally limited its focus to out-of-school variables on district NJ ASK 3 data. If out of school variables are found to explain significant variance, and in some cases predictive power in district test scores as the existing literature suggests, the value of using district test scores to measure the quality of in-school variables may be in question.

Summary of Bivariate Correlational Findings for the Dependent Variables

For this study, NJ ASK 3 Language Arts and Math school district scores were considered the dependent variables and paired with the following independent variables:

1. Median household income per district
2. Percentage of families below the poverty level
3. Percentage of economically disadvantaged students
4. Percentage of household annual income under $30,000
5. Percentage of household annual income above $200,000
6. Percentage of lone-parent households
7. Level of parental education--percentage 25 years or older, no high school diploma
8. Level of parental education--percentage 25 years or older, high school graduate
9. Level of parental education--percentage 25 years or older, high school graduate with some college experience
10. Level of parental education--percentage 25 years or older, bachelor's degree
11. Level of parental education--percentage 25 years or older, advanced degree

To determine the significance, strength, and direction of the relationship between each independent variable and NJ ASK 3 LAL and Math scores, the Pearson Correlational Coefficient for each relationship was calculated using the SPSS software.

Table 3.

Pearson Correlation Coefficients Scores for All Variables where N = 438

Correlations

		% Passing LAL	% Passing Math	% families below poverty
% Passing LAL	Pearson Correlation	1	.834**	-.582**
	Sig. (2-tailed)		.000	.000
	N	438	438	438
% Passing Math	Pearson Correlation	.834**	1	-.524**
	Sig. (2-tailed)	.000		.000
	N	438	438	438
% Families below poverty	Pearson Correlation	-.582**	-.524**	1
	Sig. (2-tailed)	.000	.000	
	N	438	438	438
% $200,000 incme	Pearson Correlation	.622**	.496**	-.386**
	Sig. (2-tailed)	.000	.000	.000
	N	438	438	438
% less $30,000 income	Pearson Correlation	-.646**	-.574**	.834**
	Sig. (2-tailed)	.000	.000	.000
	N	438	438	438
% lone-parent	Pearson Correlation	-.642**	-.563**	.751**
	Sig. (2-tailed)	.000	.000	.000
	N	438	438	438
% advanced degree	Pearson Correlation	.559**	.456**	-.230**
	Sig. (2-tailed)	.000	.000	.000
	N	438	438	438
% bachelor's degree	Pearson Correlation	.671**	.568**	-.519**
	Sig. (2-tailed)	.000	.000	.000
	N	438	438	438

% HS Diip some college	Pearson Correlation	-.633**	-.568**	.637**
	Sig. (2-tailed)	.000	.000	.000
	N	438	438	438
% HS Dip	Pearson Correlation	-.584**	-.493**	.322**
	Sig. (2-tailed)	.000	.000	.000
	N	438	438	438
% No HS Dip	Pearson Correlation	-.636**	-.527**	.685**
	Sig. (2-tailed)	.000	.000	.000
	N	438	438	438
% ED	Pearson Correlation	-.368**	-.318**	.359**
	Sig. (2-tailed)	.000	.000	.000
	N	438	438	438
Median income	Pearson Correlation	.211**	.177**	-.239**
	Sig. (2-tailed)	.000	.000	.000
	N	438	438	438

Correlations

		% 200,000 incme	% less $30,000 income	% lone parent	%adv degree
% Passing LAL	Pearson Correlation	.622**	-.646**	-.642**	.559**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% Passing Math	Pearson Correlation	.496**	-.574**	-.563**	.456**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% families below poverty	Pearson Correlation	-.386**	.834**	.751**	-.230**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% 200,000 incme	Pearson Correlation	1	-.595**	-.554**	.834**
	Sig. (2-tailed)		.000	.000	.000
	N	438	438	438	438
% Less $30,000 income	Pearson Correlation	-.595**	1	.829**	-.376**
	Sig. (2-tailed)	.000		.000	.000
	N	438	438	438	438
% Lone-parent	Pearson Correlation	-.554**	.829**	1	-.396**

	Sig. (2-tailed)	.000	.000		.000
	N	438	438	438	438
% Adv degree	Pearson Correlation	.834**	-.376**	-.396**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	438	438	438	438
% Bach degree	Pearson Correlation	.776**	-.675**	-.599**	.755**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% HS Dip some college	Pearson Correlation	-.597**	.660**	.797**	-.555**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% HS Dip	Pearson Correlation	-.794**	.524**	.456**	-.836**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% No HS Dip	Pearson Correlation	-.569**	.716**	.639**	-.529**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% ED	Pearson Correlation	-.298**	.382**	.391**	-.227**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
Median income	Pearson Correlation	.338**	-.323**	-.272**	.246**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438

Correlations

		% bach degree	% HS Dip some college	% HS Dip	% No HS Dip
% Passing LAL	Pearson Correlation	.671**	-.633**	-.584**	-.636**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% Passing Math	Pearson Correlation	.568**	-.568**	-.493**	-.527**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% families below poverty	Pearson Correlation	-.519**	.637**	.322**	.685**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% $200,000 incme	Pearson Correlation	.776**	-.597**	-.794**	-.569**

	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% Less $30,000 income	Pearson Correlation	-.675**	.660**	.524**	.716**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% Lone-parent	Pearson Correlation	-.599**	.797**	.456**	.639**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% AAdv degree	Pearson Correlation	.755**	-.555**	-.836**	-.529**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
% Bach degree	Pearson Correlation	1	-.644**	-.880**	-.748**
	Sig. (2-tailed)		.000	.000	.000
	N	438	438	438	438
% HS Dip some college	Pearson Correlation	-.644**	1	.493**	.503**
	Sig. (2-tailed)	.000		.000	.000
	N	438	438	438	438
% HS Dip	Pearson Correlation	-.880**	.493**	1	.512**
	Sig. (2-tailed)	.000	.000		.000
	N	438	438	438	438
% No HS Dip	Pearson Correlation	-.748**	.503**	.512**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	438	438	438	438
% ED	Pearson Correlation	-.303**	.339**	.223**	.379**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438
median income	Pearson Correlation	.275**	-.278**	-.290**	-.208**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	438	438	438	438

Correlations

		% ED	median income
% Passing LAL	Pearson Correlation	-.368**	.211**

	Sig. (2-tailed)	.000	.000
	N	438	438
% Passing Math	Pearson Correlation	-.318**	.177**
	Sig. (2-tailed)	.000	.000
	N	438	438
% families below poverty	Pearson Correlation	.359**	-.239**
	Sig. (2-tailed)	.000	.000
	N	438	438
% $200,000 incme	Pearson Correlation	-.298**	.338**
	Sig. (2-tailed)	.000	.000
	N	438	438
% Less $30,000 income	Pearson Correlation	.382**	-.323**
	Sig. (2-tailed)	.000	.000
	N	438	438
% Lone-parent	Pearson Correlation	.391**	-.272**
	Sig. (2-tailed)	.000	.000
	N	438	438
% Adv degree	Pearson Correlation	-.227**	.246**
	Sig. (2-tailed)	.000	.000
	N	438	438
% Bach degree	Pearson Correlation	-.303**	.275**
	Sig. (2-tailed)	.000	.000
	N	438	438
% HS Dip some college	Pearson Correlation	.339**	-.278**
	Sig. (2-tailed)	.000	.000
	N	438	438
% HS Dip	Pearson Correlation	.223**	-.290**
	Sig. (2-tailed)	.000	.000
	N	438	438
% No HS Dip	Pearson Correlation	.379**	-.208**
	Sig. (2-tailed)	.000	.000
	N	438	438
% ED	Pearson Correlation	1	-.361**
	Sig. (2-tailed)		.000
	N	438	438
Median income	Pearson Correlation	-.361**	1

	Sig. (2-tailed)	.000	
	N	438	438

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

Interpretation of Pearson Correlation Coefficients for Dependent Variable: NJ ASK 3 LAL scores

The Pearson Correlation Coefficients measure the degree of association between each variable. The correlation coefficient is not a proportion. The correlation coefficient values range from -1.00 to 1.00. To interpret correlation coefficient values the following scale was applied: .8 and above = strong, .6 - .8 = moderate strong, .4 - .6 = moderate, .2 - .4 = weak, 0 - .2 = little, if any. In Table 3 the correlation coefficients are listed from strongest to weakest. Positive and negative signs are ignored when determining the strength of coefficients. A positive value implies a positive association, whereas a large independent variable tends to be associated with a larger dependent variable. Conversely, a negative association implies a larger independent variable tends to be associated with a smaller dependent variable. The significance for all of the pairings was determined to be 0.000. This indicated all the relationship of all the predictors to be significant with a very low probability that the relationships are randomly associated.

The percentage of households in each school district with a bachelor's degree proved to have a moderate strong association with 2009 NJ ASK 3 LAL scores. The percentage of families with less than $30,000 annual income also proved to have a moderate strong association with 2009 NJ ASK 3 LAL scores.

The percentage of economically disadvantaged families proved to have a weak association with 2009 NJ ASK 3 LAL scores. Median Income also proved to have a weak association with 2009 NJ ASK 3 LAL scores.

Table 4.

Pearson Correlation Coefficients for 2009 NJ ASK3 LAL scores where N = 438

Variable	Pearson Correlation Coefficient
% Bachelor's degree	.671
% Families less than $30,000 annual income	-.646
% Lone-parent household	-.642
% No high school diploma	-.636
% High school diploma, some college	-.633
% Families more than $200,000 annual income	.622
% High school diploma	-.584
% Families below poverty	-.582
% Advanced degree	.559
% Economically disadvantaged	-.368
Median income	.211

Note. * = p<.05, **=p<.01, ***p<.001

Interpretation of Pearson Correlation Coefficients for Dependent Variable: NJ ASK 3 Math scores

The Pearson Correlation Coefficients measure the degree of association between each independent variable and the dependent variable. The correlation coefficient is not a proportion. The correlation coefficient values range from -1.00 to 1.00. To interpret correlation coefficient values the following scale was applied: .8 and above = strong, .6 - .8 = moderate strong, .4 - .6 = moderate, .2 - .4 = weak, 0 - .2 = little, if any. In Table 4 the correlation coefficients are listed from strongest to weakest. Positive and negative signs are ignored when determining strength of coefficients. A positive value implies a positive association, whereas a large independent variable tends to be associated with a larger dependent variable. Conversely, a negative association implies a larger

independent variable tends to be associated with a smaller dependent variable. The significance for all of the pairings was determined to be 0.000. This indicated all the relationships of all the predictors to be significant with a very low probability the relationships are randomly associated.

The percentage of families in each school district with less than $30,000 annual income proved to have a moderate association with 2009 NJ ASK 3 Math scores. The percentage of households with a bachelor's degree also proved to have a moderate association with 2009 NJ ASK 3 Math scores. These are the same two variables found to have the strongest association with 2009 NJ ASK 3 LAL scores.

The percentage of economically disadvantaged families proved to have a weak association with 2009 NJ ASK 3 Math scores. Median income also proved to have a weak association with 2009 NJ ASK 3 Math scores. These are the same two variables found to have the weakest association with 2009 NJ ASK 3 LAL scores.

Table 5.

Pearson Correlation Coefficients for 2009 NJ ASK3 Math scores where N = 438

Predictor Variable	**Pearson Correlation Coefficient**
% Families less than $30,000 annual income	-.574
% Bachelor's degree	.568
% High school diploma some college	-.568
% Lone-parent household	-.563
% No high school diploma	-.527
% Families below poverty	-.524
% Families more than $200,000 annual income	.496
% High school diploma	-.493

% Advanced degree	.456
% Economically disadvantaged	-.318
Median income	.177

Note: * = p<.05, **=p<.01, ***p<.001

Results of Stepwise Multiple Regression of Dependent Variable: LAL

After careful examination of the Pearson Correlation Coefficient for each predictor variable and determining the significance of each relationship, a stepwise linear multiple regression analysis of the predictor variables and dependent variable was conducted. Five models were created.

Interpretation of Stepwise Multiple Regression Model Summary for 2009 NJ ASK 3 LAL scores

The stepwise multiple regression estimates the impact of five models on 2009 NJ ASK 3 LAL scores (dependent variable). For Model 1 the predictor percentage of population with a bachelor's degree reports an R Square of .450 and explains 45% of the variance in the dependent variable. In Model 2 the predictor percentage of lone-parent households is added and reports an R Square of .538. Therefore, Model 2 demonstrates the combination of predictors: percentage of population with a bachelor's degree and predictor of percentage of lone-parent households explains 53% of the variance in the dependent variable. In Model 3 the predictor percentage of population with an advanced degree is added and reports an R Square of .553. Therefore, Model 3 demonstrates the combination of predictors: percentage of population with a bachelor's degree, percentage of lone parent-households and percentage of population with an advanced degree explains 55% of the variance in the dependent variable. In Model 4 the predictor

percentage of families below poverty is added and reports an R Square of .574. Model 4 demonstrates the combination of predictors: percentage of population with a bachelor's degree, percentage of lone parent household, percentage of population with an advanced degree and percentage of families below poverty explains 57% of the variance in the dependent variable. In Model 5 the predictor percentage of economically disadvantaged families is added and reports an R Square of .579. Model 5 demonstrates the combination of predictors: percentage of population with a bachelor's degree, percentage of lone-parent households, percentage of population with an advanced degree, percentage of families below poverty and percentage of economically disadvantaged families explains 58% of the variance in the dependent variable. Of the five models, Model 5 explains the greatest amount of variance in the dependent variable.

Table 6.

Model Summary of Stepwise Multiple Regression Model for 2009 NJ ASK3 LAL

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.671[a]	.450	.449	11.19641
2	.735[b]	.540	.538	10.25428
3	.743[c]	.553	.550	10.12439
4	.757[d]	.574	.570	9.89583
5	.761[e]	.579	.575	9.84075

[a] Predictors: (Constant), % bach degree. [b] Predictors: (Constant), % bach degree, % lone-parent. [c] Predictors:(Constant), % bach degree, % lone-parent, % adv degree. [d] Predictors: (Constant), % bach degree, % lone parent, %adv degree, % families below poverty. [e] Predictors: (Constant), % bach degree, % lone parent, %adv degree, % families below poverty, % ED

Interpretation of Two-Way ANOVA for Stepwise Multiple Regression Model for 2009 NJ ASK3 LAL Scores

This two-way ANOVA estimates the impact of five main effects on the dependent variable in five different models. The ANOVA demonstrates all five models are statistically significant.

Model 1 is significant at the .000 level, F = 357.432, df =1, 436.

Model 2 is significant at the .000 level, F = 255.463, df = 2, 435.

Model 3 is significant at the .000 level, F = 178,784, df = 3, 434.

Model 4 is significant at the .000 level, F = 145,674, df = 4, 433.

Model 5 is significant at the .000 level, F = 119,019, df = 5, 43Table 7.

Table 7

Two-Way ANOVA Stepwise Multiple Regression Model for 2009 NJ ASK LAL Scores

ANOVA[f]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	44807.570	1	44807.570	357.432	.000[a]
	Residual	54656.768	436	125.360		
	Total	99464.338	437			
2	Regression	53724.008	2	26862.004	255.463	.000[b]
	Residual	45740.331	435	105.150		
	Total	99464.338	437			
3	Regression	54977.922	3	18325.974	178.784	.000[c]
	Residual	44486.416	434	102.503		
	Total	99464.338	437			
4	Regression	57061.749	4	14265.437	145.674	.000[d]
	Residual	42402.589	433	97.927		

	Total	99464.338	437			
5	Regression	57629.316	5	11525.863	119.019	.000[e]
	Residual	41835.022	432	96.840		
	Total	99464.338	437			

[a] Predictors: (Constant), % bach degree. [b] Predictors: (Constant), % bach degree, % lone-parent. [c] Predictors: (Constant), % bach degree, % lone-parent, % adv degree. [d] Predictors: (Constant), % bach degree, % lone-parent, % adv degree, % families below poverty. [e]Predictors: (Constant), % bach degree, % lone-parent, % adv degree, % families below poverty, % ED [f]Dependent Variable: % Passing LAL.

Interpretation of Standardized Coefficient Betas and Tolerance for Stepwise Multiple Regression Model for 2009 NJ ASK3 LAL Scores

The coefficient table demonstrates how each predictor influences the dependent variable. In Model 1 the predictor percentage of population with a bachelor's degree reports a beta = .671. It is statistically significant at the .000 level, t = 18.906. The beta is positive, which means as the percentage of population with a bachelor's degree increases, 2009 NJ ASK 3 LAL scores increase. In Model 2, the predictor percentage of population with a bachelor's degree decreases in power with a beta = .447. It is significant at the .000 level, t = 11.018. The predictor added in Model 2, the percentage of lone-parent households, reports a beta = -.374. It is significant at the .000 level, t = 9.209. The negative beta for the percentage of lone-parent households indicates that as percentage of lone-parent households increases, 2009 NJ ASK 3 LAL scores decrease.

In Model 3, the predictor percentage of population with a bachelor's degree decreases in power again with a beta = .308. It is significant at the .000 level, t = 5.469. The predictor percentage of lone-parent households gains some power with a beta = -.389. It is significant at the .000 level, t = 9.646. The predictor added in Model 3, percentage of population with an advanced degree, reports a beta = .172. It is significant

at the .001 level, t = 3.498. The predictor percentage of population with an advanced degree was the least power of the three predictors in Model 3; however, its inclusion did increase the power of the percentage of lone-parent households.

In Model 4, the predictor percentage of population with a bachelor's degree decreases further in power with a beta = .229. It is significant at the .000 level, t = 3.964. The predictor percentage of lone-parent households decreased in power with a beta = -.236. It is significant at the .000 level, t = 4.595. The predictor percentage of population with an advanced degree increased its power with a beta = .239. It is significant at the .000 level, t= 4.764. The predictor added to Model 4, percentage of families below poverty, reports a beta = -.231. It is significant at the .000 level, t = 4.613. The negative beta for percentages of families below poverty indicates that as the percentage of families below poverty increase, 2009 NJ ASK 3 LAL scores decrease.

In Model 5, the predictor percentage of population with a bachelor's degree decreased slightly further in power with a beta = .227. It is significant at the .000 level, t = 3.995. The predictor percentage of lone-parent households decreased in power slightly with a beta = -.218. It is significant at the .000 level, t = 4.210. The predictor percentage of population with an advanced degree decreased its power slightly with a beta = .232. It is significant at the .000 level, t= 4.641. The predictor percentage of families below poverty decreased its power slightly with a beta = -.218. It is significant at the .000 level, t = 4.345. The predictor added to Model 5, the percentage of economically disadvantaged families, reports a beta = -.083. It is significant at the .016 level, t = 2.421. In Model 5, the predictor with the greatest power is the percentage of population with an

advanced degree. Interestingly, this predictor gained power in both Model 4 and Model 5 when two predictors related to household income were added to the models.

The VIF for all predictors in all the models fell below the threshold of 5, which would indicate the models do not have a multicollinearity problem. However, some of the models reported significantly lower VIF numbers than other models. The model with more than one predictor and the lowest VIF rating was Model 2. In Model 2, both predictors reported a VIF of 1.559. Model 2 was the only model in which all predictors, except for Model 1, in which only one predictor was present, were less than 2. Therefore, Model 2 can be considered the model with the least multicollinearity.

Table 8.

Standardized Coefficient Betas and Tolerance for Stepwise Multiple Regression Model for 2009 NJ ASK3 LAL Scores

Coefficients[a]

Model		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		Beta			Tolerance	VIF
1	(Constant)		27.518	.000		
	% Bach degree	.671	18.906	.000	1.000	1.000
2	(Constant)		24.585	.000		
	% Bach degree	.447	11.018	.000	.641	1.559
	% Lone-parent	-.374	-9.209	.000	.641	1.559
3	(Constant)		25.129	.000		
	% Bach degree	.308	5.469	.000	.324	3.088
	% Lone-parent	-.389	-9.646	.000	.634	1.577
	% Adv degree	.172	3.498	.001	.426	2.349
4	(Constant)		25.923	.000		
	% Bach degree	.229	3.964	.000	.295	3.389
	% Lone-parent	-.236	-4.595	.000	.372	2.687
	% Adv degree	.239	4.764	.000	.390	2.565

	% families below poverty	-.231	-4.613	.000	.393	2.547
5	(Constant)		26.172	.000		
	% Bach degree	.227	3.955	.000	.295	3.390
	% Lone-parent	-.218	-4.210	.000	.364	2.747
	% Adv degree	.232	4.641	.000	.388	2.574
	% families below poverty	-.218	-4.345	.000	.388	2.578
	% ED	-.083	-2.421	.016	.829	1.207

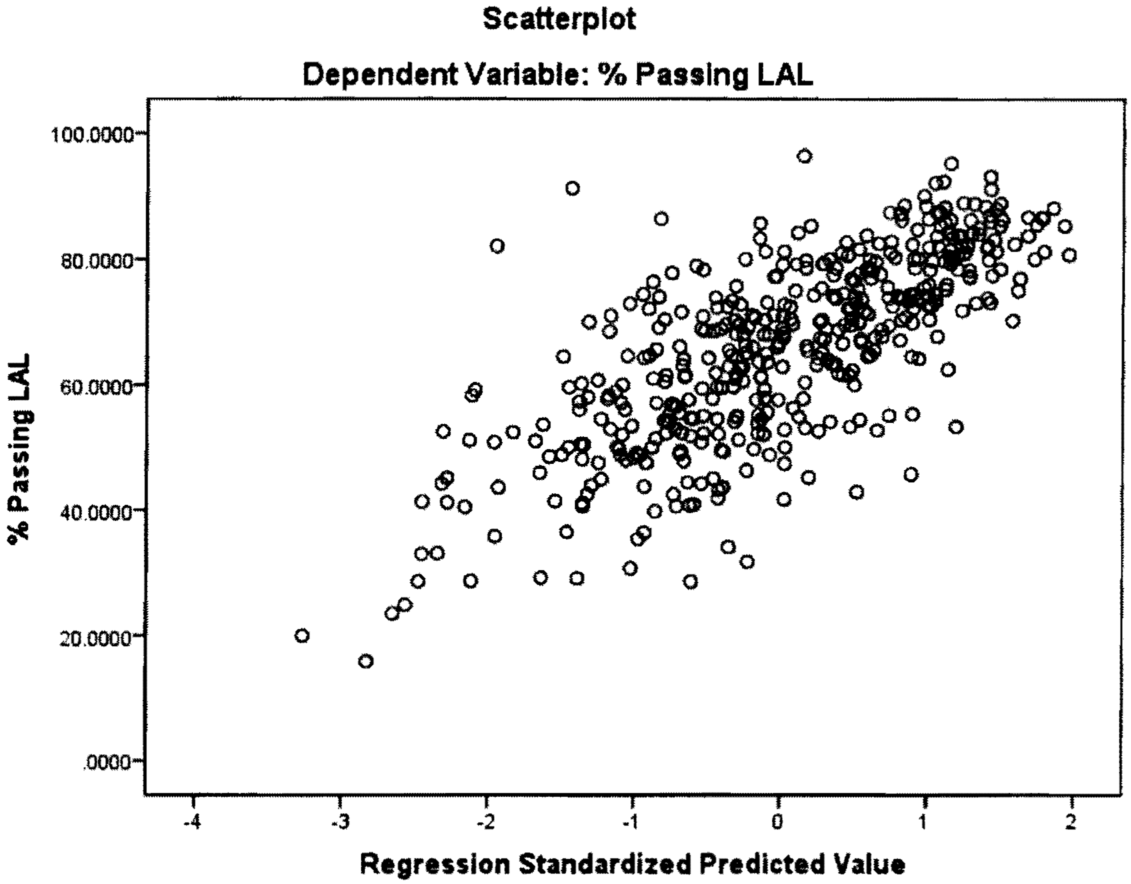
[a] Dependent Variable: % Passing LAL

Results of Theoretical Framework Multiple Regression of Dependent Variable:LAL

The extant review of literature suggested a model different from the stepwise models might be best. Rather than rely solely on the stepwise regression models, theoretical support for the results was also given consideration. Consequently, various simultaneous regression models were run based on the evidence found in the extant literature. This evidence suggested that the variables of household income, the percentage of lone-parent households, and the level of parental education within a school district may combine to predict student achievement as measured by standardized tests.

Table 9

Scatterplot Theoretical Framework Multiple Regression Predictive Formula for NJ ASK 3 LAL Scores



Interpretation of Model Summary for Theoretical Framework Multiple Regression for 2009 NJ ASK 3 LAL Model

The theoretical framework multiple regression models were constructed based on the three constructs of theoretical framework: household income, lone-parent household, and parent education. Each construct included at least one predictor variable.

All different variations of the one predictor variable from each construct were tested to determine which combination explained the greatest variance in the dependent variable with a VIF for all three predictors less than 2. The VIF standard was established to ensure the model didn't suffer from multicollinearity. For Model 1, the predictor percentage of lone-parent household, percentage of economically disadvantaged families,

and percentage of population with a bachelor's degree reports an R Square of .549 and explains 54% of the variance in the dependent variable.

Table 10.

Model Summary of Theoretical Framework Multiple Regression Model

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.741[a]	.549	.546	10.16695

[a] Predictors: (Constant), % lone parent, % ED, % bach degree

Interpretation of Two-Way ANOVA for Theoretical Framework Multiple Regression for 2009 NJ ASK 3 LAL Scores

This two-way ANOVA estimates the impact of three main effects on the dependent variable in one model. The ANOVA demonstrates the model is significant at the .000 level, F = 176.082, df =3, 434.

Table 11.

Two-Way ANOVA for Theoretical Framework Multiple Regression for 2009 NJ ASK LAL Model

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	54603.156	3	18201.052	176.082	.000[a]
	Residual	44861.182	434	103.367		
	Total	99464.338	437			

[a] Predictors: (Constant), % lone parent, % ED, % bach degree [b] Dependent Variable: % Passing LAL

Interpretation of Standardized Coefficient Betas and Tolerance for Theoretical Framework Multiple Regression for 2009 NJ ASK 3 LAL Model

The coefficient table demonstrates how each predictor influences the dependent variable. In Model 1 the predictor percentage of economically disadvantaged families reports a beta = -.103. It is statistically significant at the .004 level, t = 2.916. The beta is negative, which means as the percentage of economically disadvantaged families increases, 2009 NJ ASK 3 LAL scores decrease. The second predictor percentage of population with a bachelor degree's degree reports a beta = .436. It is significant at the .000 level, t = 10.789. The third predictor percentage of lone-parent households, reports a beta = -.340. It is significant at the .000 level, t = 8.131. The negative beta for percentage of lone-parent households indicates that as the percentage of lone-parent households increases, 2009 NJ ASK 3 LAL scores decrease. All three predictors in Model 1 reported a VIF of less than 2 and tolerance levels all exceed 1- R^2 (tolerances exceed .451), demonstrating multicollinearity is within acceptable limits.

Table 12.

Standardized Coefficient Betas & Tolerance for Theoretical Framework Multiple Regression for 2009 NJ ASK3 LAL Model

Coefficients[a]

Model		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		Beta			Tolerance	VIF
1	(Constant)		24.966	.000		
	% ED	-.103	-2.916	.004	.840	1.191
	% bach degree	.436	10.789	.000	.636	1.573
	% lone parent	-.340	-8.131	.000	.593	1.686

[a] Dependent Variable: % Passing LAL

Results of Hierarchical Linear Regression of Dependent Variable: LAL

Hierarchical linear regression of the three independent variables identified by the Theoretical Framework was conducted to further explain the predictive relationship on the dependent variable. Hierarchical linear regression is a more advanced form of simple linear regression and can provide further insight into how the independent variables impact one another and the dependent variable. For the purpose of this hierarchical linear regression, the independent variable from the theoretical framework with the highest beta was entered first. The independent variable with the second highest beta was entered second and the independent variable with the lowest beta was entered last. Three different models were generated.

Interpretation of Hierarchical Linear Regression for 2009 NJ ASK 3 LAL Scores

The hierarchical linear regression estimates the impact of three models on 2009 NJ ASK 3 LAL scores (dependent variable). For Model 1 the predictor percentage of population with a bachelor's degree reports an R Square of .450 and explains 45% of the variance in the dependent variable. In Model 2 the predictor percentage of lone-parent households is added and reports an R Square of .540. Therefore, Model 2 demonstrates the combination of predictors: percentage of population with a bachelor's degree and predictor of percentage of lone-parent households explains 54% of the variance in the dependent variable. In Model 3 the predictor percentage of economically disadvantaged families is added and reports an R Square of .549. Therefore, Model 3 demonstrates the combination of predictors: percentage of population with a bachelor's degree, percentage of lone-parent households, and percentage of economically disadvantaged families

explains 55% of the variance in the dependent variable. Of the three models, Model 3 explains the greatest amount of variance in the dependent variable.

Table 13.

Model Summary of Hierarchical Linear Regression Model for 2009 NJ ASK 3 LAL

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.671[a]	.450	.449	11.1964083
2	.735[b]	.540	.538	10.2542764
3	.741[c]	.549	.546	10.1669454

Model Summary

Model	Change Statistics				
	R Square Change	F Change	df1	df2	Sig. F Change
1	.450	357.432	1	436	.000
2	.090	84.797	1	435	.000
3	.009	8.505	1	434	.004

[a] Predictors: (Constant), % bach degree. [b] Predictors: (Constant), % bach degree, % lone-parent. [c] Predictors: (Constant), % bach degree, % lone-parent, % ED

Interpretation of Two-Way ANOVA for Hierarchical Linear Regression Model for 2009 NJ ASK 3 LAL Scores

This two-way ANOVA estimates the impact of three main effects on the dependent variable in three different models. The ANOVA demonstrates all three models are statistically significant.

Model 1 is significant at the .000 level, F = 357.432, df =1, 436.

Model 2 is significant at the .000 level, F = 255.463, df = 2, 435.

Model 3 is significant at the .000 level, F = 176.082, df = 3, 434.

Table 14.

Two-Way ANOVA Hierarchical Linear Regression Model for 2009 NJ ASK LAL scores

		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	44807.571	1	44807.571	357.432	.000[a]
	Residual	54656.768	436	125.360		
	Total	99464.338	437			
2	Regression	53724.008	2	26862.004	255.463	.000[b]
	Residual	45740.331	435	105.150		
	Total	99464.338	437			
3	Regression	54603.156	3	18201.052	176.082	.000[c]
	Residual	44861.182	434	103.367		
	Total	99464.338	437			

[a] Predictors: (Constant), % bach degree. [b] Predictors: (Constant), % bach degree, % lone parent
[c] Predictors: (Constant), % bach degree, % lone parent, % ED[d] Dependent Variable: % Passing LAL

Interpretation of Standardized Coefficient Betas and Tolerance for Hierarchical Linear Regression Model for 2009 NJ ASK3 LAL Scores

The coefficient table demonstrates how each predictor influences the dependent variable. In Model 1 the predictor percentage of population with a bachelor's degree reports a beta = .671. It is statistically significant at the .000 level, t = 18.906. The beta is positive, which means as the percentage of population with a bachelor's degree increases, 2009 NJ ASK 3 LAL scores increase. In Model 2, the predictor percentage of population with a bachelor's degree decreases in power with a beta = .447. It is significant at the .000 level, t = 11.018. The predictor added in Model 2, percentage of lone-parent households, reports a beta = -.374. It is significant at the .000 level, t =

9.209. The negative beta for percentage of lone-parent households indicates that as the percentage of lone-parent households increases, 2009 NJ ASK 3 LAL scores decrease.

In Model 3, the predictor percentage of population with a bachelor's degree decreases in power again with a beta = .436. It is significant at the .000 level, t = 10.789. The predictor percentage of lone-parent households loses some power with a beta = -.340. It is significant at the .000 level, t = 8.131. The predictor added in Model 3, the percentage of economically disadvantaged families, reports a beta = -.103. It is significant at the .004 level, t = 2.916. The predictor percentage of economically disadvantaged families was the least powerful of the three predictors in Model 3.

The VIF for all predictors in all the models fell below the threshold of 5 and the tolerances all exceed 1- R^2 (tolerances exceed .451), which would indicate the models do not have a multicollinearity problem. The results of the hierarchical linear regression are very similar to the simultaneous multiple regression of the theoretical framework model. This provides more evidence of the predictive reliability of these independent variables on the dependent variable.

Table 15.

Standardized Coefficient Betas & Tolerance for Hierarchical Linear Regression Model for 2009 NJ ASK 3 LAL scores

Coefficients[a]

Model		Standardized Coefficients	t	Sig.	Correlations		
		Beta			Zero-order	Partial	Part
1	(Constant)		27.518	.000			
	% Bach degree	.671	18.906	.000	.671	.671	.671
2	(Constant)		24.585	.000			
	% Bach degree	.447	11.018	.000	.671	.467	.358

	% Lone-parent	-.374	-9.209	.000	-.642	-.404	-.299
3	(Constant)		24.966	.000			
	% Bach degree	.436	10.789	.000	.671	.460	.348
	% Lone-parent	-.340	-8.131	.000	-.642	-.364	-.262
	% ED	-.103	-2.916	.004	-.368	-.139	-.094

Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	(Constant)		
	% Bach degree	1.000	1.000
2	(Constant)		
	% Bach degree	.641	1.559
	% Lone-parent	.641	1.559
3	(Constant)		
	% Bach degree	.636	1.573
	% Lone-parent	.593	1.686
	% ED	.840	1.191

[a] Dependent Variable: % Passing LAL

Examples of Predictive Power for Dependent Variable: LAL

In total, six statistically significant models were identified through multiple regressions with different combinations of predictors to explain the variance in the dependent variable. To determine which model produced the strongest predictive power the betas and constant for each model were entered into the following formula from Maylone (2002).

$$A_i (X_i) + A_{ii} (X_{ii}) + A_{iii} (X_{iii}) \ldots + \text{Constant} = Y$$

A_i = individual school district predictor value

X_i = beta for predictor

Y = predicted LAL score

A predicted score was calculated from each of the six models for the entire population. These scores were entered into the database for the dependent variable and then the margin of error was calculated by subtracting the predicted score from the actual score for the entire population. Last, the standard deviation for each distribution of margin of error data was calculated. The model with the lowest standard deviation for distribution of margin of error was concluded to have the greatest predictive power.

Table 16.

Standard Deviation for Distribution of Margin of Error for All Six Multiple Regression Models for 2009 NJ ASK 3 LAL Scores where N = 438

Model	Standard Deviation	Rank
A (theoretical framework)	10.53	1
1	11.94	6
2	10.70	2
3	10.79	3
4	11.19	5
5	10.98	4

Example 1: Maple Shade Twp (DFG: CD)

For the Maple Shade Township school district, the values for the three-out-of school variables (% lone-parent household, % bachelor's degree, % economically disadvantaged) are as follows:

a = % lone parent household = 37.51

b = % bachelor degree = 14.63

c = % economically disadvantaged = 38.00

Enter these values into the following equation:

$$-.034(37.51) + .436(14.63) + -0.103(38) + 59.004 = 48.71$$

The result, 48.71, represents the predicted LAL score for Maple Shade Township school district 2009 NJ ASK 3. It suggests 48.71 % of Grade 3 students enrolled at Maple Shade Township are predicted to score either Proficient or Advanced Proficient. The actual percentage of Grade 3 students enrolled at Maple Shade Township in 2009 that scored either Proficient or Advanced Proficient on LAL NJ ASK 3 equaled 48.70%. The difference for the predicted score was calculated by subtracting the actual score from the predicted score. For example: 48.70 – 48.71 = -0.01.

Example 2: Vineland City (DFG: A)

For the Vineland City school district, the values for the three out-of-school variables (% lone-parent household, % bachelor's degree, % economically disadvantaged) are as follows:

a = % lone-parent household	=	42.71
b = % bachelor's degree	=	11.09
c = % economically disadvantaged	=	10.00

Enter these values into the following equation:

$$-.034(42.71) + .436(11.09) + -0.103(10) + 59.004 = 48.29$$

The result, 48.29, represents the predicted LAL score for Vineland City school district 2009 NJ ASK 3. It suggests 48.29 % of Grade 3 students enrolled at Vineland City are predicted to score either Proficient or Advanced Proficient. The actual percentage of Grade 3 students enrolled at Vineland City in 2009 that scored either Proficient or Advanced Proficient on LAL NJ ASK 3 equaled 48.10%. The difference for the

predicted score was calculated by subtracting the actual score from the predicted score. For example: 48.10 – 48.29 = -0.19.

Example 3: Harding Township (DFG: J)

For the Harding Township school district, the values for the three out-of-school variables (% lone-parent household, % bachelor'S degree, % economically disadvantaged) are as follows:

a = % lone parent-household	=	16.73
b = % bachelor's degree	=	37.42
c = % economically disadvantaged	=	0.00

Enter these values into the following equation:

$$-.034(16.73) + .436(37.42) + -0.103(0.00) + 59.004 = 69.93$$

The result, 69.93, represents the predicted LAL score for Harding Township school district 2009 NJ ASK 3. It suggests 69.95 % of Grade 3 students enrolled at Harding Township are predicted to score either Proficient or Advanced Proficient. The actual percentage of Grade 3 students enrolled at Harding Township in 2009 that scored either Proficient or Advanced Proficient on LAL NJ ASK 3 equaled 73%. The difference for the predicted score was calculated by subtracting the actual score from the predicted score. For example: 73.00 – 69.93 = 3.37.

Example 4: Mount Arlington (DFG: GH)

For the Mount Arlington school district, the values for the three out-of-school variables (% lone-parent household, % bachelor'S degree, % economically disadvantaged) are as follows:

a = % lone-parent household	=	20.90

b = % bachelor's degree = 28.19

c = % economically disadvantaged = 27.00

Enter these values into the following equation:

$$-.034(20.90) + .436(28.19) + -0.103(27.00) + 59.004 = 61.40$$

The result, 61.40, represents the predicted LAL score for Mount Arlington school district 2009 NJ ASK 3. It suggests 61.40 % of Grade 3 students enrolled at Mount Arlington are predicted to score either Proficient or Advanced Proficient. The actual percentage of Grade 3 students enrolled at Mount Arlington in 2009 that scored either Proficient or Advanced Proficient on LAL NJ ASK 3 equaled 61.50%. The difference for the predicted score was calculated by subtracting the actual score from the predicted score. For example: 61.50 – 61.40 = .10.

Summary of Analysis for Dependent Variable: LAL

The percentage of households in each school district with a bachelor's degree proved to have a moderate strong association with 2009 NJ ASK 3 LAL scores. The percentage of families with less than $30,000 annual income also proved to have a moderate strong association with 2009 NJ ASKS 3 LAL scores.

The percentage of economically disadvantaged families proved to have a weak association with 2009 NJ ASK3 LAL scores. Median income also proved to have a weak association with 2009 NJ ASK3 LAL scores.

In the stepwise regression, Model 5 demonstrated the combination of predictors: percentage of population with a bachelor's degree, percentage of lone-parent household, percentage of population with an advanced degree, percentage of families below poverty, and percentage of economically disadvantaged families explains 58% of the variance in

the dependent variable. Of all the stepwise regression models, Model 5 explains the greatest amount of variance in the dependent variable.

Within the theoretical construct for parental education, four of the six independent variables were shown to have an inverse relationship with 2009 NJ ASK3 LAL scores. Both the percentage of population with a bachelor's degree and the percentage of the population with an advanced degree were shown to have a positive relationship with 2009 NJ ASK 3 LAL scores. The independent variable shown to explain the least amount of variance in 2009 NJ ASK 3 LAL scores was median household income within a school district. Median household income was shown to have an inverse relationship with 2009 NJ ASK 3 LAL scores and explain 21.1 percent of the variance within the scores.

Of the five stepwise regression models and the one theoretical framework simultaneous multiple regression model, the theoretical framework identified the best predictive model for 2009 NJ ASK 3 LAL scores. The standard deviation of the margin of error for the predicted scores was the smallest of all the models, at 10.53. This model predicted 52% of the population's 2009 NJ ASK 3 LAL scores with a +/- 10 margin of error. In other words, 227 of 438 predicted scores were within 10 points of the actual scores. The theoretical framework model explained 54.9% of the variance on 2009 NJ ASK3 LAL scores. This model included the following three independent variables: percentage of lone-parent household, percentage bachelor's degree. and percentage economically disadvantaged. Maylone's (2002) predictive formula explained 56.1% of the variance in the dependent variable.

Hierarchical linear regression was then applied to the theoretical framework model to further test its predictive reliability. The results of this regression were almost identical to the simultaneous multiple regression with no multicollinearity problems, which provides greater evidence for the predictive power of these three independent variables.

Results of Stepwise Multiple Regression of Dependent Variable: Math

After careful examination of the Pearson Correlation Coefficient for each predictor variable and determining the significance of each relationship, a stepwise linear multiple regression analysis of the predictor variables and dependent variable was conducted. Six models were created.

Interpretation of Stepwise Multiple Regression Model Summary for 2009 NJ ASK 3 Math scores

The stepwise multiple regression estimates the impact of six models on 2009 NJ ASK 3 Math scores (dependent variable). For Model 1 the predictor percentage of families with less than $30,000 annual income reports an R Square of .330 and explains 33% of the variance in the dependent variable. In Model 2 the predictor percentage of families with less than $30,000 annual income and percentage of population with an advanced degree reports an R Square of .397. Therefore, Model 2 demonstrates the combination of the predictors: percentage of families with less than $30,000 annual income and percentage of population with an advanced degree explains 39.7% of the variance in the dependent variable. In Model 3 the predictor percentage of population with high school diploma and some college is added and reports an R Square of .420. Therefore Model 3 demonstrates the combination of predictors: percentage of families

with less than \$30,000 annual income, percentage of population with an advanced degree and percentage of population with a high school diploma and some college explains 42% of the variance in the dependent variable. In Model 4 the predictor percentage of population with no high school diploma is added and reports an R Square of .429. Therefore, Model 4 demonstrates the combination of predictors: percentage of families with less than \$30,000 annual income, percentage of population with an advanced degree, percentage of population with a high school diploma and some college, and percentage of population with no high school diploma explains 42.9% of the variance in the dependent variable. In Model 5 the predictor percentage of population with a high school diploma is added and reports an R Square of .436. Therefore, Model 5 demonstrates the combination of predictors: percentage of families with less than \$30,000 annual income, percentage of population with an advanced degree, percentage of population with a high school diploma and some college, percentage of population with no high school diploma and percentage of population with a high school diploma explains 43.6% of the variance in the dependent variable. In Model 6 the predictor percentage of population with an advanced degree is removed and reports an R Square of .436. Therefore, Model 6 demonstrates the combination of predictors: percentage of families with less than \$30,000 annual income, percentage of population with a high school diploma and some college, percentage of population with no high school diploma and percentage of population with a high school diploma explains 43.6% of the variance in the dependent variable. Of the six models, Model 6 explains the greatest amount of variance in the dependent variable with the fewest predictors.

Table 17.

Model Summary of Stepwise Multiple Regression Model for 2009 NJ ASK 3 Math Scores

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.574[a]	.330	.328	10.5956
2	.630[b]	.397	.394	10.0610
3	.648[c]	.420	.416	9.8754
4	.655[d]	.429	.423	9.8159
5	.661[e]	.436	.430	9.7604
6	.661[f]	.436	.431	9.7494

[a] Predictors: (Constant), % less $30,000 income. [b] Predictors: (Constant), % less $30,000 income, % adv degree. [c] Predictors: (Constant), % less $30,000 income, % adv degree, % HS Dip some college. [d] Predictors: (Constant), % less $30,000 income, % adv degree, % HS Dip some college, % No HS Dip. [e] Predictors: (Constant), % less $30,000 income, % adv degree, % HS Dip some college, % No HS Dip, % HS Dip. [f] Predictors: (Constant), % less $30,000 income, % HS Dip some college, % No HS Dip, % HS Dip

Interpretation of Two-Way ANOVA for Stepwise Multiple Regression Model for 2009 NJ ASK 3 Math Scores

This two-way ANOVA estimates the impact of six main effects on the dependent variable in six different models. The ANOVA demonstrates all six models are significant.

Model 1 is significant at the .000 level, F = 214.518, df =1, 436.

Model 2 is significant at the .000 level, F = 143.239, df = 2, 435.

Model 3 is significant at the .000 level, F = 104,950, df = 3, 434.

Model 4 is significant at the .000 level, F = 81.242, df = 4, 433.

Model 5 is significant at the .000 level, F = 86.920, df = 5, 432.

Model 6 is significant at the .000 level, F = 83.832, df = 4, 432.

Table 18.

Two-Way ANOVA for Stepwise Multiple Regression 2009 NJ ASK Math Models

ANOVA[g]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	24083.040	1	24083.040	214.518	.000[a]
	Residual	48947.989	436	112.266		
	Total	73031.029	437			
2	Regression	28998.551	2	14499.275	143.239	.000[b]
	Residual	44032.478	435	101.224		
	Total	73031.029	437			
3	Regression	30705.505	3	10235.168	104.950	.000[c]
	Residual	42325.524	434	97.524		
	Total	73031.029	437			
4	Regression	31310.914	4	7827.729	81.242	.000[d]
	Residual	41720.115	433	96.351		
	Total	73031.029	437			
5	Regression	31876.073	5	6375.215	66.920	.000[e]
	Residual	41154.956	432	95.266		
	Total	73031.029	437			
6	Regression	31873.646	4	7968.412	83.832	.000[f]
	Residual	41157.383	433	95.052		
	Total	73031.029	437			

[a] Predictors: (Constant), % less $30,000 income. [b] Predictors: (Constant), % less $30,000 income, % adv degree. [c] Predictors: (Constant), % less $30,000 income, % adv degree, % HS Dip some college. [d] Predictors: (Constant), % less $30,000 income, % adv degree, % HS Dip some college, % No HS Dip. [e] Predictors: (Constant), % less $30,000 income, % adv degree, % HS Dip some college, % No HS Dip, % HS Dip. [f] Predictors: (Constant), % less $30,000 income, % HS Dip some college, % No HS Dip, % HS Dip. [g] Dependent Variable: % Passing Math

Interpretation of Standardized Coefficient Betas and Tolerance for Stepwise Multiple Regression Model for 2009 NJ ASK3 Math scores

The coefficient table demonstrates how each predictor influences the dependent variable. In Model 1 predictor percentage of families with less than $30,000 annual income reports a beta = -.574. It is statistically significant at the .000 level, t = 14.646. The beta is negative, which means as the percentage of families with less than $30,000 annual income increases, 2009 NJ ASK 3 Math scores decrease. In Model 2, the predictor percentage of families with less than $30,000 annual income decreases in power with a beta = -.469. It is significant at the .000 level, t = 11.671. The predictor added in Model 2, percentage of population with an advanced degree reports a beta = .280. It is significant at the .000 level, t = 6.969. The positive beta for percentage of population with an advanced degree indicates that as the percentage of population with an advanced degree increases, 2009 NJ ASK 3 LAL scores increase. Within Model 2 the percentage of families with less than $30,000 annual income is almost twice as powerful as the predictor percentage of population with an advanced degree.

In Model 3, the predictor percentage of families with less than $30,000 annual income decreases in power again with a beta = -.350. It is significant at the .000 level, t = 7.190. The predictor percentage of population with an advanced degree loses power with a beta = .199. It is significant at the .000 level, t = 4.533. The predictor added in Model 3, percentage of population with high school diploma and some college, reports a beta = -.227. It is significant at the .000 level, t = 4.184. The predictor percentage of population with an advanced degree was the least power of the three predictors in Model 3.

In Model 4, the percentage of families with less than $30,000 annual income decreases further in power with a beta = .253. It is significant at the .000 level, t = 4.083.

The predictor percentage of population with an advanced degree decreased in power with a beta = .148. It is significant at the .002 level, t = 3.086. The predictor percentage of population with a high school diploma and some college increased its power slightly with a beta = -.246. It is significant at the .000 level, t= 4.527. The predictor added to Model 4, percentage of population with no high school diploma, reports a beta = -.144. It is significant at the .013 level, t = 2.507. Of the four predictors in Model 4, percentage of population with high school diploma and some college gained power to be almost equal with percentage of families with less than $30,000 annual income.

In Model 5, the predictor percentage of families with less than $30,000 annual income decreased further in power with a beta = -.169. It was significant at the .017 level, t = 2.389. The predictor percentage of population with an advanced degree decreased its power further with a beta = -.013 and reversed its direction. It also proved to have a multicollinearity problem with a VIF = 5.125. The predictor percentage of population with a high school diploma and some college gained in power with a beta = -.284. It is significant at the .000 level, t = 5.046. The predictor percentage of population with no high school diploma increased its power slightly with a beta = -.176. It is significant at the .003 level, t= 2.999. The predictor added to Model 5, percentage of population with a high school diploma, reports a beta = -.186. It is significant at the .015 level, t = 2.436. In Model 5, the predictor percentage of population with an advanced degree demonstrated a multicollinearity problem, which disqualifies Model 5 from consideration.

In Model 6, the predictor percentage of population with advanced degrees was removed. The predictor percentage of families with less than $30,000 annual income

gained some power with a beta = -.175. It was significant at the .004 level, t = 2.908. The predictor percentage of population with a high school diploma and some college reported a beta = -.280. It was significant at the .000 level, t = 5.666. The predictor percentage of population with no high school diploma reported a beta = -.172. It was significant at the .001 level, t = 3.223. The predictor percentage of population with a high school diploma reported a beta = -.176. It was significant at .000 level, t = 3.947. All four predictors in Model 6 demonstrated a negative direction in their relationship with the dependent variable.

The VIF for all predictors in all the models except Model 5 fell below the threshold of 5 and the tolerances exceed 1- R^2, which would indicate the models do not have a multicollinearity problem. However, some of the models reported significantly lower VIF numbers than other models. The model with more than one predictor and the lowest VIF rating was Model 2. In Model 2, both predictors reported a VIF of 1.165. Model 2 was the only model in which all predictors, except for Model 1 in which only one predictor was present, demonstrated a VIF less than 2 and tolerances exceed 1- R^2. Therefore, Model 2 can be considered to be the model with the least multicollinearity.

Table 19.

Standardized Coefficient Betas and Tolerance for Stepwise Multiple Regression 2009 NJ ASK 3 Math Models

Coefficients[a]

Model		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		Beta			Tolerance	VIF
1	(Constant)		83.374	.000		
	% less $30,000 income	-.574	-14.646	.000	1.000	1.000

2	(Constant)		48.539	.000		
	% Less $30,000 income	-.469	-11.671	.000	.859	1.165
	% Adv degree	.280	6.969	.000	.859	1.165
3	(Constant)		35.797	.000		
	% Less $30,000 income	-.350	-7.190	.000	.564	1.772
	% Adv degree	.199	4.533	.000	.692	1.445
	% HS Dip some college	-.227	-4.184	.000	.455	2.198
4	(Constant)		34.817	.000		
	% Less $30,000 income	-.253	-4.083	.000	.344	2.906
	% Adv degree	.148	3.086	.002	.570	1.754
	% HS Dip some college	-.246	-4.527	.000	.446	2.245
	% No HS Dip	-.144	-2.507	.013	.400	2.498
5	(Constant)		19.324	.000		
	% Less $30,000 income	-.169	-2.389	.017	.262	3.822
	% Adv degree	-.013	-.160	.873	.195	5.125
	% HS Dip some college	-.284	-5.046	.000	.412	2.427
	% No HS Dip	-.176	-2.999	.003	.381	2.627
	% HS Dip	-.186	-2.436	.015	.224	4.474
6	(Constant)		57.551	.000		
	% Less $30,000 income	-.175	-2.908	.004	.361	2.768
	% HS Dip some college	-.280	-5.666	.000	.535	1.871
	% No HS Dip	-.172	-3.233	.001	.462	2.165
	% HS Dip	-.176	-3.947	.000	.653	1.531

[a] Dependent Variable: % Passing Math

Results of Theoretical Framework Multiple Regression of Dependent Variable: Math

The extant review of literature suggested its own model, different from the stepwise models. Rather than rely solely on the stepwise regression models, theoretical support for the results was also given consideration. Consequently, various simultaneous regression models were run based on the evidence found in the extant literature. This

evidence suggested that the variable of household income, the percentage of lone-parent household, and the level of parental education within a school district may combine to predict student achievement as measured by standardized tests.

Interpretation of Model Summary for Theoretical Framework Multiple Regression for 2009 NJ ASK 3 Math Model

The theoretical framework multiple regression models were constructed based on the three constructs of theoretical framework: household income, lone-parent household and parental education. Each construct included at least one predictor variable. All different variations of the one predictor variable from each construct were tested to determine which combination explained the greatest variance in the dependent variable with a VIF for all three predictors equaling less than 2. The VIF standard was established to ensure the model did not suffer from multicollinearity. For Model 1 the predictor percentage of lone-parent household, percentage of economically disadvantaged families, and percentage of population with a bachelor's degree reports an R Square of .406 and explains 40.6% of the variance in the dependent variable.

Table 20.

Scatterplot Theoretical Framework Multiple Regression Predictive Formula for 2009 NJ ASK 3 Math scores

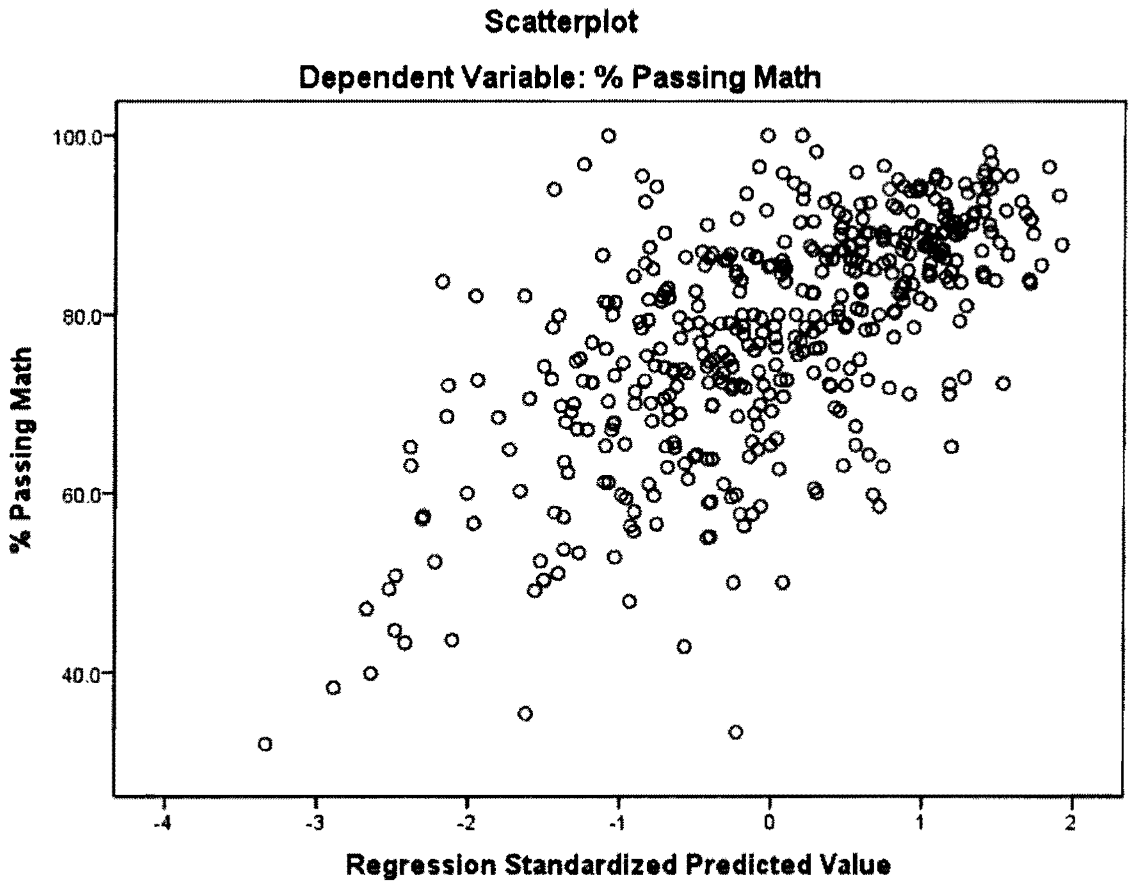


Table 21.

Model Summary of Theoretical Framework Multiple Regression Model for 2009 NJ ASK 3 Math

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.637[a]	.406	.402	9.9970

[a] Predictors: (Constant), % lone-parent, % ED, % bach degree

Interpretation of Two-Way ANOVA for Theoretical Framework Multiple Regression for 2009 NJ ASK 3 Math Scores

This two-way ANOVA estimates the impact of three main effects on the dependent variable in one model. The ANOVA demonstrates the model is significant at the .000 level, F = 98.915, df =3, 434.

Table 22.

Two-Way ANOVA for Theoretical Framework Multiple Regression 2009 NJ ASK 3 Math Model

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	29656.830	3	9885.610	98.915	.000[a]
	Residual	43374.199	434	99.941		
	Total	73031.029	437			

[a] Predictors: (Constant), % lone parent, % ED, % bach degree
[b] Dependent Variable: % Passing Math

Interpretation of Standardized Coefficient Betas and Tolerance for Theoretical Framework Multiple Regression for 2009 NJ ASK 3 Math Model

The coefficient table demonstrates how each predictor influences the dependent variable. In Model 1 the predictor percentage of economically disadvantaged families reports a beta = -.086. It is statistically significant at the .033 level, t = 2.140. The beta is negative, which means as the percentage of economically disadvantaged families increases, 2009 NJ ASK3 Math scores decrease. The second predictor percentage of population with a bachelor degree reports a beta = .350. It is significant at the .000 level, t = 7.555. The third predictor percentage of lone-parent households reports a beta = -.319. It is significant at the .000 level, t = 6.644. The negative beta for percentage of

lone-parent households indicates that as the percentage of lone-parent households increases, 2009 NJ ASK 3 Math scores decrease. All three predictors in Model 1 reported a VIF of less than 2 and the tolerances exceeded $1 - R^2$ (tolerances greater than .594), all within acceptable ranges demonstrating a low likelihood of multicollinearity problem.

Table 23.

Standardized Coefficient Betas & Tolerance for Theoretical Framework Multiple Regression 2009 NJ ASK 3 Math Model

Coefficients[a]

Model		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		Beta			Tolerance	VIF
1	(Constant)		31.908	.000		
	% ED	-.086	-2.140	.033	.840	1.191
	% bach degree	.350	7.555	.000	.636	1.573
	% lone parent	-.319	-6.644	.000	.593	1.686

[a] Dependent Variable: % Passing Math

Results of Hierarchical Multiple Regression of Dependent Variable: Math

Hierarchical linear regression of the three independent variables indentified by the theoretical framework was conducted to further explain their predictive relationship to the dependent variable. Hierarchical linear regression is a more advanced form of simple linear regression and can provide further insight into how the independent variables impact one another and the dependent variable. For the purpose of this hierarchical linear regression, the independent variable from the theoretical framework with the highest beta was entered first. The independent variable with the second highest beta was entered

second and the independent variable with the lowest beta was entered last. Three different models were generated.

Interpretation of Hierarchical Linear Regression for 2009 NJ ASK 3 Math Scores

The hierarchical linear regression estimates the impact of three models on 2009 NJ ASK3 Math scores (dependent variable). For Model 1 the predictor percentage of population with a bachelor's degree reports an R Square of .322 and explains 32% of the variance in the dependent variable. In Model 2 the predictor percentage of lone-parent household is added and reports an R Square of .40. Therefore, Model 2 demonstrates the combination of predictors: percentage of population with a bachelor's degree and predictor of percentage of lone-parent household explains 40% of the variance in the dependent variable. In Model 3 the predictor percentage of economically disadvantaged families is added and reports an R Square of .406. Therefore, Model 3 demonstrates the combination of predictors: percentage of population with a bachelor's degree, percentage of lone-parent household and percentage of economically disadvantaged families explains 41% of the variance in the dependent variable. Of the three models, Model 3 explains the greatest amount of variance in the dependent variable.

Table 24.

Model Summary of Hierarchical Multiple Linear Regression Model for 2009 NJ ASK 3 Math

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.568[a]	.322	.321	10.6533
2	.632[b]	.400	.397	10.0381

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.568[a]	.322	.321	10.6533
2	.632[b]	.400	.397	10.0381
3	.637[c]	.406	.402	9.9970

Model Summary

Model	Change Statistics				
	R Square Change	F Change	df1	df2	Sig. F Change
1	.322	207.481	1	436	.000
2	.077	56.087	1	435	.000
3	.006	4.579	1	434	.033

[a] Predictors: (Constant), % bach degree. [b] Predictors: (Constant), % bach degree, % lone parent.[c] Predictors: (Constant), % bach degree, % lone parent, % ED.

Interpretation of Two-Way ANOVA for Hierarchical Linear Regression Model for 2009 NJ ASK 3 Math Scores

This two-way ANOVA estimates the impact of three main effects on the dependent variable in three different models. The ANOVA demonstrates all three models are statistically significant.

Model 1 is significant at the .000 level, F = 207.481, df =1, 436.

Model 2 is significant at the .000 level, F = 144.891, df = 2, 435.

Model 3 is significant at the .000 level, F = 98.915, df = 3, 4

Table 25.

Two-Way ANOVA Hierarchical Linear Regression Model for 2009 NJ ASK 3 Math Scores

ANOVA[d]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	23547.768	1	23547.768	207.481	.000[a]
	Residual	49483.261	436	113.494		

	Total	73031.029	437			
2	Regression	29199.209	2	14599.604	144.891	.000[b]
	Residual	43831.820	435	100.763		
	Total	73031.029	437			
3	Regression	29656.830	3	9885.610	98.915	.000[c]
	Residual	43374.199	434	99.941		
	Total	73031.029	437			

[a] Predictors: (Constant), % bach degree, [b] Predictors: (Constant), % bach degree, % lone parent.
[c] Predictors: (Constant), % bach degree, % lone parent, % ED.
[d] Dependent Variable: % Passing Math

Interpretation of Standardized Coefficient Betas & Tolerance for Hierarchical Linear Regression Model for 2009 NJ ASK 3 LAL Scores

The coefficient table demonstrates how each predictor influences the dependent variable. In Model 1 the predictor percentage of population with a bachelor's degree reports a beta = .568. It is statistically significant at the .000 level, t = 14.404. The beta is positive, which means as the percentage of population with a bachelor's degree increases, 2009 NJ ASK 3 LAL scores increase. In Model 2, the predictor percentage of population with a bachelor's degree decreases in power with a beta = .360. It is significant at the .000 level, t = 7.758. The predictor added in Model 2, percentage of lone-parent households, reports a beta = -.347. It is significant at the .000 level, t = 7.489. The negative beta for percentage of lone-parent household indicates that as the percentage of lone parent-household increases, 2009 NJ ASK 3 LAL scores decrease.

In Model 3, the predictor percentage of population with a bachelor's degree decreases in power again with a beta = .350. It is significant at the .000 level, t = 7.555. The predictor percentage of lone-parent households loses some power with a beta = -

.319. It is significant at the .000 level, t = 6.644. The predictor added in Model 3, percentage of economically disadvantaged families, reports a beta = -.086. It is significant at the .033 level, t = 2.140. The predictor percentage of economically disadvantaged families was the least powerful of the three predictors in Model 3.

The VIF for all predictors in all the models fell below the threshold of 5 and the tolerances exceeded $1 - R^2$ (tolerances greater than .594), which would indicate that the models do not have a multicollinearity problem. The results of the hierarchical linear regression are almost identical to the simultaneous multiple regression of the theoretical framework model. This provides more evidence of the predictive reliability of these independent variables on the dependent variable.

Table 26.

Standardized Coefficient Betas & Tolerance for Hierarchical Linear Regression Model for 2009 NJ ASK 3 Math Scores

Coefficients[a]

Model		Standardized Coefficients	t	Sig.	Correlations		
		Beta			Zero-order	Partial	Part
1	(Constant)		42.632	.000			
	% Bach degree	.568	14.404	.000	.568	.568	.568
2	(Constant)		31.761	.000			
	% Bach degree	.360	7.758	.000	.568	.349	.288
	% Lone-parent	-.347	-7.489	.000	-.563	-.338	-.278
3	(Constant)		31.908	.000			
	% Bach degree	.350	7.555	.000	.568	.341	.279
	% Lone-parent	-.319	-6.644	.000	-.563	-.304	-.246
	% ED	-.086	-2.140	.033	-.318	-.102	-.079

Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	(Constant)		
	% Bach degree	1.000	1.000
2	(Constant)		
	% Bach degree	.641	1.559
	% Lone-parent	.641	1.559
3	(Constant)		
	% Bach degree	.636	1.573
	% Lone-parent	.593	1.686
	% ED	.840	1.191

[a] Dependent Variable: % Passing Math

Examples of Predictive Power for Dependent Variable: Math

In total, seven statistically significant models were identified through multiple regressions with different combinations of predictors to explain the variance in the dependent variable. To determine which model produced the strongest predictive power, the betas and constant for each model were entered into the below formula from Maylone (2002).

$$A_i\,(X_i\,) + A_{ii}\,(X_{ii}) + A_{iii}\,(X_{iii}) \ldots + \text{Constant} = Y$$

A_i = individual school district predictor value

X_i = beta for predictor

Y = predicted Math score

A predicted score was calculated from each of the seven models for the entire population. These scores were entered into the database for the dependent variable and

then the margin of error was calculated by subtracting the predicted score from the actual score for the entire population. Lastly, the standard deviation for each distribution of margin of error data was calculated. The model with the lowest standard deviation for distribution of margin of error was concluded to have the greatest predictive power.

Table 27.

Standard Deviation for Distribution of Margin of Error for All Six Multiple Regression Models for2009 NJ ASK 3 Math Scores where N = 438

Model	Standard Deviation	Rank
B (theoretical framework)	10.07	1
1	10.79	6
2	10.58	5
3	10.56	4
4	11.00	7
5	10.32	2
6	10.33	3

Example 1:Maple Shade Twp (DFG: CD)

For the Maple Shade Township school district, the values for the three out-of-school variables (% lone-parent household, % bachelor'S degree, % economically disadvantaged) were as follows:

a = % lone-parent household = 37.51

b = % bachelor's degree = 14.63

c = % economically disadvantaged = 38.00

Entering these values into the following equation:

$$-0.0319(37.51) + .350(14.63) + -0.086(38) + 74.15 = 64.04$$

The result, 64.04 represents the predicted Math score for Maple Shade Township school district 2009 NJ ASK 3. It suggests 64.04 % of Grade 3 students enrolled at Maple Shade Township are predicted to score either Proficient or Advanced Proficient. The actual percentage of Grade 3 students enrolled at Maple Shade Township in 2009 that scored either Proficient or Advanced Proficient on Math NJ ASK 3 equaled 61.20%. The margin of error for the predicted score was calculated by subtracting the actual score from the predicted score. For example, 61.20 – 64.04 = -2.84.

Example 2: Vineland City (DFG: A)

For the Vineland City school district, the values for the three out-of-school variables (% lone-parent household, % bachelor'S degree, % economically disadvantaged) were as follows:

a = % lone-parent household	=	42.71
b = % bachelor's degree	=	11.09
c = % economically disadvantaged	=	10.00

Enter these values into the following equation:

$$-0.319(42.71) + .350(11.09) + -0.086(10) + 74.15 = 63.55$$

The result, 63.55, represents the predicted Math score for the Vineland City school district 2009 NJ ASK 3. It suggests 63.55 % of Grade 3 students enrolled at Vineland City are predicted to score either Proficient or Advanced Proficient. The actual percentage of Grade 3 students enrolled at Vineland City in 2009 that scored either Proficient or Advanced Proficient on Math NJ ASK 3 equaled 63.5%. The margin of error for the predicted score was calculated by subtracting the actual score from the predicted score. For example, 63.5 – 63.55 = -0.05.

Example 3: Harding Township (DFG: J)

For the Harding Township school district, the values for the three out of school variables (% lone parent household, % bachelor degree, % economically disadvantaged):

a = % lone parent household = 16.73

b = % bachelor degree = 37.42

c = % economically disadvantaged = 0.00

Enter these values into the following equation:

$$-0.319(16.73) + .350(37.42) + -0.086(0.00) + 74.15 = 81.91$$

The result, 81.91, represents the predicted Math score for Harding Township school district 2009 NJ ASK 3. It suggests 81.91% of Grade 3 students enrolled at Harding Township are predicted to score either Proficient or Advanced Proficient. The actual percentage of Grade 3 students enrolled at Harding Township in 2009 that scored either Proficient or Advanced Proficient on Math NJ ASK 3 equaled 73%. The margin of error for the predicted score was calculated by subtracting the actual score from the predicted score. For example, 73.00 – 81.91 = - 8.91.

Example 4: Mount Arlington (DFG: GH)

For the Mount Arlington school district, the values for the three out-of-school variables (% lone-parent household, % bachelor'S degree, % economically disadvantaged) were as follows:

a = % lone-parent household = 20.90

b = % bachelor's degree = 28.19

c = % economically disadvantaged = 27.00

Enter these values into the following equation:

$$-0.319(20.90) + .350(28.19) + -0.086(27.00) + 74.15 = 75.02$$

The results, 75.02, represents the predicted Math score for Mount Arlington school district 2009 NJ ASK 3. It suggests 75.02 % of Grade 3 students enrolled at Mount Arlington are predicted to score either Proficient or Advanced Proficient. The actual percentage of Grade 3 students enrolled at Mount Arlington in 2009 that scored either Proficient or Advanced Proficient on Math NJ ASK 3 equaled 74.4%. The margin of error for the predicted score was calculated by subtracting the actual score from the predicted score. For example, 74.4 – 75.02 = -0.62.

Summary of Analysis for Dependent Variable: Math

The percentage of families in each school district with less than $30,000 annual income proved to have a moderate association with 2009 NJ ASK 3 Math scores. The percentage of households with a bachelor's degree also proved to have a moderate association with 2009 NJ ASK3 Math scores. These are the same two variables found to have the strongest association with 2009 NJ ASK 3 LAL scores.

The percentage of economically disadvantaged families proved to have a weak association with 2009 NJ ASK 3 Math scores. Median income also proved to have a weak association with 2009 NJ ASK 3 Math scores. These are the same two variables found to have the weakest association with 2009 NJ ASK3 LAL scores.

In the stepwise regression Model 6, the predictor percentage of population with an advanced degree is removed and reports an R Square of .436. Therefore, Model 6 demonstrates the combination of predictors: percentage of families with less than $30,000 annual income, percentage of population with high school diploma and some college, percentage of population with no high school diploma, and percentage of population with

high school diploma explains 43.6% of the variance in the dependent variable. Of the six models, Model 6 explains the greatest amount of variance in the dependent variable with the fewest predictors.

As with the results for 2009 NJ ASK 3 LAL scores, when the theoretical construct was applied to 2009 NJ ASK 3 Math, it demonstrated a similar impact relative to parental education. The same four of the six independent variables were shown to have an inverse relationship with 2009 NJ ASK3 LAL scores. Both the percentage of population with a bachelor's degree and the percentage of the population with an advanced degree were shown to have a positive relationship with 2009 NJ ASK 3 LAL scores. The independent variable shown to explain the least amount of variance in 2009 NJ ASK3 Math scores was median household income within a school district. The median household income was shown to have an inverse relationship with 2009 NJ ASK3 Math scores and explain 17.7 percent of the variance within the scores.

As with 2009 NJ ASK 3 LAL scores, the theoretical framework simultaneous multiple regressions identified the predictive model of 2009 NJ ASK 3 Math scores with the smallest standard deviation in the margin of error, at 10.07. Furthermore, this model predicted 60% of the population 2009 NJ ASK 3 Math scores with a less than +/- 10 margin of error. In other words, 263 of 438 predicted scores were less than 10 points from the actual scores. The theoretical framework model explained 40.6% of the variance on 2009 NJ ASK 3 Math scores. This model included the same three independent variables as the model for 2009 NJ ASK 3 LAL scores. They are percentage of lone-parent households, percentage of bachelor's degree, and percentage of

economically disadvantaged. Maylone's (2002) predictive formula explained 56.1% of the variance in the dependent variable.

Hierarchical linear regression was then applied to the theoretical framework model to further test its predictive reliability. The results of this regression were very similar to the simultaneous multiple regression with no multicollinearity problems, which provides greater evidence for the predictive power of these three independent variables.

Research Questions and Answers for Dependent Variables

This study began by examining three overarching research questions: (1) How much variance in NJ ASK 3 test results in Language Arts and Mathematics is explained by out-of-school socioeconomic variables? (2) How accurately can out-of-school socio-economic and community-level variables predict a school district's percentage of students scoring Proficient or above on the NJ ASK 3 Language Arts and Mathematics sections? and (3) Which community-level variables account for the greatest amount of variance in a school district's percentage of students passing the NJ ASK 3?

To gain a deeper understanding about these questions and after a thorough review of extant literature, four research questions were developed.

Research Question 1: How much variance in 2009 NJ ASK 3 test results in Language Arts can be explained by the household income construct (Table 2) for New Jersey school districts?

Null Hypothesis 1: No statistically significant relationship exists between 2009 NJ ASK 3 scores in Language Arts and the household income construct for New Jersey school districts.

Answer: The null hypothesis is rejected. All of the predictor variables from the household income construct were statistically significant predictors of 2009 NJ ASK 3 scores in Language Arts.

Research Question 2: How much variance in 2009 NJ ASK 3 test results in Mathematics can be explained by the household income construct (Table 2) for New Jersey school districts?

Null Hypothesis 2: No statistically significant relationship exists between 2009 NJ ASK 3 scores in Mathematics and the household income construct for New Jersey school districts.

Answer: The null hypothesis is rejected. All of the predictor variables from the household income construct were statistically significant predictors of 2009 NJ ASK 3 scores in Mathematics.

Research Question 3: How much variance in 2009 NJ ASK 3 test results in Language Arts can be explained by the lone-parent household construct (Table 3) for New Jersey school districts?

Null Hypothesis 3: No statistically significant relationship exists between 2009 NJ ASK 3 test results in Language Arts and the lone-parent household construct for New Jersey school districts.

Answer: The null hypothesis is rejected. The predictor variables from the lone-parent household construct were statistically significant predictors of 2009 NJ ASK 3 scores in Language Arts.

Research Question 4: How much variance in 2009 NJ ASK 3 test results in Mathematics can be explained by the lone-parent household construct (Table 3) for New Jersey school districts?

Null Hypothesis 4: No statistically significant relationship exists between 2009 NJ ASK 3 test results in Mathematics and the lone-parent household construct for New Jersey school districts.

Answer: The null hypothesis is rejected. The predictor variables from the lone - parent household construct were statistically significant predictors of 2009 NJ ASK 3 scores in Mathematics.

Research Question 5: How much variance in 2010 NJ ASK 3 test results in Language Arts can be explained by the parental education construct (Table 4) for New Jersey school districts?

Null Hypothesis 5: No statistically significant relationship exits between 2009 NJ ASK 3 test results in Language Arts and the level of parental education construct for New Jersey school districts.

Answer: The null hypothesis is rejected. All of the predictor variables from the parental education construct were statistically significant predictors of 2009 NJ ASK 3 scores in Language Arts.

Research Question 6: How much variance in 2010 NJ ASK 3 test results in Mathematics can be explained by the parental education construct (Table 4) for New Jersey school districts?

Null Hypothesis 6: No statistically significant relationship exits between 2009 NJ ASK 3 test results in Mathematics and the level of parental education construct for New Jersey school districts.

Answer: The null hypothesis is rejected. All of the predictor variables from the parental education construct were statistically significant predictors of 2009 NJ ASK 3 scores in Mathematics.

Research Question 7: Which combination of independent variables establishes the greatest reliable predictive power for a school district's 2009 NJ ASK 3 Language Arts test results?

Null Hypothesis 7: There is no statistically, research demonstrated, combination of independent variables with reliable predictive power for 2009 NJ ASK 3 test results in Language Arts and for New Jersey school districts.

Answer: The null hypothesis is rejected. A combination of percentage of lone parent household, percentage of population with a bachelor degree and percentage of economically disadvantage families was found to have reliable predictive power for 2009 NJ ASK 3 test results in Language Arts for New Jersey school districts.

Research Question 8: Which combination of independent variables establishes the greatest reliable predictive power for a school district's 2009 NJ ASK3 Mathematics test results?

Null Hypothesis 8: There is no statistically, research-demonstrated, combination of independent variables with reliable predictive power for 2009 NJ ASK 3 test results in Mathematics for New Jersey school districts.

Answer: The null hypothesis is rejected. A combination of percentage of lone-parent households, percentage of population with a bachelor's degree, and percentage of economically disadvantaged families was found to have reliable predictive power for 2009 NJ ASK 3 test results in Mathematics for New Jersey school districts.

Chapter Summary

The data for this study was analyzed by creating one database for each dependent variable. The total population for the study included 438 school districts. Simultaneous multiple linear regression and stepwise regression models were then created by importing each dependent variable database into the IBM SPSS (Statistical Package for the Social Sciences) predictive analytics software. A two-way ANOVA (analysis of variance) was generated for each dependent variable. The F-static was analyzed to determine if each regression model was statistically significant. To determine which model explained the greatest variance in each dependent variable, an analysis of each model's Adjusted R^2 (coefficient of determination) was conducted. Within each model the independent variables reported a standardized beta coefficient, which was used to compare the strength of the effect of each independent variable on the dependent variable within each statistically significant model.

One threat to the reliability and validity of the linear regression models was the impact of multicollinearity on the independent variables. While multicollinearity does not impact the overall predictive power of a regression model, it can cause individual coefficient estimates to change erratically. This can negatively impact calculations regarding the predictive power of individual school districts. Since a major aspect of this

study included the application of the formula created by Maylone (2002) to individual school districts, multicollinearity was given serious consideration.

To verify that multicollinearity did not threaten the predictive reliability of the multiple regression models generated for each dependent variable, two different methods were used to create the regression models. First, stepwise regression was applied to each dependent variable and all of the independent variables. This method sequences variables based on F-tests and Tolerance levels to build models with the greatest R^2 values and lowest multicollinearity levels. Second, the theoretical framework established through review of extant literature was applied to build simultaneous multiple regression models for each dependent variable. The condition for the application of the theoretical framework to predictive variables was the following:

- One variable from each construct must be used (household income, lone-parent household, level of parental education)
- VIF for all three variables must be less than 2.

The model which produced the highest R^2 while meeting the above two conditions was identified as the best theoretical framework model.

Of the five stepwise regression models and the one theoretical framework simultaneous multiple regression model, the theoretical framework identified the best predictive model for 2009 NJ ASK 3 LAL scores. The standard deviation of the margin of error for the predicted scores was the smallest of all the models, at 10.53. This model predicted 52% of the population's 2009 NJ ASK3 LAL scores with a +/- 10 margin of error. In other words, 227 of 438 predicted scored were within 10 points of the actual scores. The theoretical framework model explained 54.9% of the variance on 2009 NJ

ASK 3 LAL scores. This model included the following three independent variables: percentage of lone-parent households, percentage with bachelor's degree, and percentage of economically disadvantaged families. Maylone's (2002) predictive formula explained 56.1% of the variance in the dependent variable.

As with 2009 NJ ASK 3 LAL scores, the theoretical framework simultaneous multiple regressions identified the best predicative model of 2009 NJ ASK 3 Math scores with the smallest standard deviation in the margin of error at 10.07. Furthermore, this model predicted 60% of the population's 2009 NJ ASK 3 Math scores with a less than +/- 10 margin of error. In other words, 263 of 438 predicted scores were less than 10 points from the actual scores. The theoretical framework model explained 40.6% of the variance on 2009 NJ ASK 3 Math scores. This model included the same three independent variables as the model for 2009 NJ ASK 3 LAL scores. They are percentage of lone-parent households, percentage with bachelor's degree, and percentage of economically disadvantaged families. Maylone's (2002) predictive formula explained 56.1% of the variance in the dependent variable.

Hierarchical linear regression was then applied to the theoretical framework model for LAL and Math scores to further test its predictive reliability. The results of this regression were almost identical to the simultaneous multiple regression with no multicollinearity problems, which provide greater evidence for the predictive power of these three independent variables.

Chapter V

CONCLUSIONS AND RECOMMENDATIONS

The purpose for this quantitative study was to investigate the influence out-of-school variables found in the extant literature on school district student achievement as measured by the 2009 NJ ASK 3 Language Arts and Mathematics assessments. By controlling for out-of-school variables shown to influence student achievement, this study aimed to apply simultaneous multiple regression analysis to produce research-based evidence to inform education policymakers and school district leadership interpretation of 2009 NJ ASK 3 scores. A dearth of empirical evidence exists regarding the predictive power of out-of-school variables at the elementary level of school districts. Therefore, this study also added empirical results to the limited body of existing literature.

The results of this study demonstrated out-of-school variables explain 54% of the variance in school district 2009 NJ ASK 3 Language Arts scores and 40% of the variance in school district 2009 NJ ASK 3 Mathematics scores. The independent variables of percentage of lone-parent households, percentage of economically disadvantaged families, and percentage of households with a bachelor's degree combined to produce the most accurate predictive formula of school district 2009 NJ ASK 3 Language Arts and Mathematics scores. These three variables predicted 52% of 2009 NJ ASK 3 Language Arts scores and 60% of 2009 NJ ASK 3 Mathematics scores within 10 points. In other words, 228 of 438 New Jersey school district NJ ASK3 LAL scores could be predicted within 10 points by relying entirely on out-of-school variables. Furthermore, 262 of 439 New Jersey school district NJ ASK 3 Math scores could be predicted within 10 points by relying entirely on out-of-school variables. The data for these variables were based on

five-year estimates calculated by the U.S. Census Bureau from data collected from the American Community Survey. This study examined five-year estimates because they provided the largest sample size; the study examined the entire population of New Jersey school districts with at least 25 students enrolled in third grade.

This study was undertaken to explain the following questions: (1) How much variance in 2009 NJ ASK 3 test results in Language Arts and Mathematics is explained by out-of-school socioeconomic variables? (2) How accurately can out-of-school socio-economic and community-level variables predict a school district's percentage of students scoring Proficient or above on 2009 NJ ASK 3 Language Arts and Mathematics sections? and (3) Which community-level variables account for the greatest amount of variance in a school district's percentage of students passing the 2009 NJ ASK 3?

The results of the study are in agreement with Maylone (2002), where 56% of the variance in district state test scores could be explained by three out-of-school social and demographic variables: percent of students eligible for free- or reduced-lunch, percent of district lone-parent households, and mean annual district household income. "The results...also reflect the findings of an Educational Research Service study that showed that poverty alone accounted for 56% of the variance among state average test scores in the NAEP-92 Trial State Assessment in mathematics. That same study showed that a stunning 89% of those variations were due to poverty and just three other out-of-school demographic factors (number of parents living at home, parents' education, and community type)" (Educational Research Service, 1994; Maylone 2002). As federal and state governments continue to devise and adopt policies with the purported purpose of improving student achievement for all students, this study explained how specific out- of-

school variables influence student achievement as defined by a high-stakes standardized assessment. While this influence has been demonstrated in earlier studies, this study further explained this influence by showing the degree to which specific variables explain the variance in a high-stakes language arts and mathematics standardized assessment at the elementary school level. This study also identified how a combination of three variables can be used to predict the actual scores of a school district's Language Arts and Mathematics high-stakes standardized assessment at the elementary school level.

NJ ASK 3 LAL Dependent Variable

Conclusions

Based on the results from this study, more than half of the variance in 2009 NJ ASK3 language arts scores can be explained by out-of-school variables. The existing empirical literature and results from this study suggest high-stakes standardized assessment scores are significantly influenced by out-of-school factors. The three factors identified by this study to have the greatest influence are household income, single-parent homes, and level of parental education. Indeed, closer examination of the parental education variables showed how the impact on the dependent variable of 2009 NJ ASK 3 test scores reversed its direction and increased its power when a bachelor's degree had been achieved in the home. All education levels below a bachelor's degree (high school diploma or high school diploma with some college) had an inverse relationship with the dependent variable. This relationship switched to a direct relationship and increased in power when a bachelor's degree had been achieved. This study further demonstrated the importance and influence of achieving a four-year college degree on student achievement.

Household income proved to be the least powerful of the three constructs of the theoretical framework. It can be concluded that household income influences student achievement but not to the same degree as parental education or having more than one parent in the home. Of the two dependent variables, student achievement in LAL proved to be influenced more by out-of -chool variables than student achievement in math. This study showed how school districts had less influence on student achievement in writing and reading than in student achievement in math at the Grade 3 level.

NJ ASK 3 Math Dependent Variable

Conclusions

Based on the results of this study, 2009 NJ ASK 3 Math scores were less influenced by out-of-school variables than 2009 NJ ASK 3 Language Arts scores. In other words, student achievement in reading and writing appears to be influenced more by out-of-school variables than student achievement in math. This finding suggests school districts may play a greater role influencing student ability to overcome out-of - school factors in regard to learning math, while schools may be less powerful in influencing student achievement in reading and writing at the elementary school level.

Value of High-Stakes Standardized Assessment Data

Experimental social science researcher Donald T. Campbell applied the concept of "Heisenberg uncertainty," whereby in quantum physics the act of measuring something changes what is being measured, to social science research. Campbell's Law states: "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor." Applied to education policy,

Campbell's Law suggests that the greater the influence of high-stakes standardized assessment data, the less likely it will accurately measure and positively impact the quality of student learning.

Furthermore, Maylone (2002) demonstrated how out-of-school variables explained more than 50% of Michigan school districts' high school achievement scores. In some cases, Maylone (2002) applied out-of-school variables to predict school districts' actual high school test scores. Whereas Essentialist theories value the importance of standardized assessments to measure student achievement against a set of standards, Maylone (2002) suggests these standardized assessments may indeed say little about student achievement as a result of specific school district influences at the high school level. The findings from this study suggest 2009 NJ ASK 3 scores may also say little about student achievement as a result of specific school district influences at the elementary school level.

> Coleman et al. reported in 1966 that the greatest influence on student academic performance was socioeconomic status (SES), followed by teacher characteristics and class size. Over 40 years after the release of the Coleman Report (1966), much of the reviewed literature continues to support the original findings of Coleman et al. After reviewing the extensive literature available regarding the potential attainment of educational equality among students, it is evident that enacting accountability policies, providing additional funding, using high-stakes consequences and the results from those tests as major indicators of student academic success, and providing an increased number of education resources to struggling schools will not, in and of themselves, lead to the

successful bridging of existing achievement gaps at the state and national testing level (Lee & Wong, 2004; Piereira, 2011).

The findings of this study also support this conclusion.

Still high-stakes standardized assessments show persistent large achievement gaps based on a variety of social and demographic factors. "The empirical problems are obvious, especially because about half of the variance in student achievement results on standardized tests are explained by out-of-school factors, things schools cannot control. Also, the sizeable standard error of measurement (e.g., margin of test-score error) inherent in individual student test scores skews proficiency categorizations. For example, in New Jersey, the margin of error on most of the state tests is about 7-10 scale points at the proficiency cut-score. This means that about 9,500 students each year are potentially mislabeled as not proficient due to imprecise test results. The test results are not as accurate as the testing companies and state education personnel would have us believe" (Tienken, 2010). This study demonstrated with high reliability how out-of-school variables influence student achievement. Since high-stakes standardized assessments are heavily influenced by out-of-school variables and inherently flawed as a measure of student achievement, their value must be questioned.

"A cursory review of reactance theory should also raise questions about the notion of relying on standardized tests as primary outcome measures for school quality. Reactance theory states that when humans are placed in situations in which they feel they cannot succeed (e.g., 100% proficiency mandate under NCLB), feel coerced, or believe the mandates are counterproductive or harmful, they will react by doing less than their best. They will withdraw, engage in practices that are contrary to research or

recommendations, sabotage the mandate, feign minimal compliance, or openly dissent and resist (Brehm & Brehem, 1981; Silvia, 2005). In essence, the carrots and sticks used in instrumental use policies have little effect, and might have negative effects, if the people believe that they are in a no-win situation" (Tienken, 2010).

Campbell's Law and Reactance Theory demonstrate the disconnect between education policy and research while also providing further explanation for the failure of such polices, evidenced by the persistent achievement gap between social and demographic groups.

Recommendations for Policy and Practice

A stronger link needs to be made between education research and education policy. The current educational policy environment is disconnected from research and dominated by the Essentialist paradigm. This study demonstrated how the Essentialist paradigm has increasingly dominated the context of educational policy with no limits in sight for the foreseeable future.

This study and the extant literature reviewed demonstrate the flaws of relying solely on high-stakes standardized assessments to measure and define student achievement. A more robust definition of student achievement, which accounts for the influence of research-demonstrated social and demographic factors, is needed. Education policy driven by a more robust definition of student achievement will be incentivized to move away from standardized measures of learning and toward more formative measures. "The existing empirical literature and the results from this study seem to suggest that the more proximal (closer to the student) the formative assessment activity is (i.e., self-evaluation), the greater the influence it has on learning" (Pereira, 2011).

Research suggests that educational policy relying on standardized assessment to define student learning will continue to produce the same results and demonstrate the same gaps in achievement based on social and demographic categories. Research suggests that standardized assessments do not significantly improve student learning.

Indeed, the results from this study raise the question: Why bother with the tests when many district test results (more than half in this study) can be predicted within 10 percentage points using out-of-school variables? At the very least, this study demonstrates how a more balanced definition of student achievement is required. While standardized assessments may serve as one indicator of student achievement, more are needed. Most importantly, the need remains to view the learner as a co-creator of knowledge, exposed to many influences which research demonstrates influence student learning. Therefore, policy must be designed to incentivize educationally sound and research-based practices consistent with Progressive theories about teaching and learning. To do so, will ensure policies are constructed from and rooted in the American tradition of individualism rather than the Essentialist values of sameness.

At the very least, policymakers and school leaders desirous of improving language arts and math test scores would be well served to allocate resources and identify strategies to combat out-of-school variables within a school district shown to influence student achievement. This fact, however, has been known for decades and some legislation to combat the conditions of poverty has been attempted. However, more aggressive polices are warranted to combat the negative effects of poverty to improve the quality of learning for all students.

The current education reform environment remains dominated by a psychometrically driven version of the Essentialist philosophy. As demonstrated in Chapter II, this dominance has grown steadily since the early 1980s despite a weak base of research. Consequently, the focus remains on standardized assessment to define student achievement. Indeed, *A Blueprint for Reform: The Reauthorization of the Elementary and Secondary Education Act* describes how the focus on standardized assessments will increase. "In essence, the *Blueprint* cements a commitment to nationalize and standardize education and it has the potential to shift the governance of public schools further from a locally controlled endeavor and closer to a centrally planned operation" (Tienken 2010).

The sad irony of Essentialist policies is that their stated intention to create equity and justice for all students by improving the ability of all students to access quality education causes the exact opposite. These policies fail to recognize the individual needs of each student and consequently incentivize behaviors, which will not result in improved test scores. These behaviors include the implementation of a highly standardized curriculum for all, an increase in tracking of students and an increase in the skimming or exclusion of students from certain populations. Essentialist policies aimed at creating social justice for the most at-risk students actually incentivize oppressive and exclusionary practices because they do not account for the individual needs of the learner. Essentialist theory fails to recognize one size does not fit all.

The theoretical framework for this study established a research base for a potential new policy context to (1) determine which school districts are improving student learning while controlling for specific out-of-school variables, (2) re-

contextualize standard assessment data from high-stakes to representing one aspect of a larger student, school, and district profile, (3) redefine student achievement where demonstration of proficiency can be shown in multiple ways and recognizes students develop in stages, and (4) provide empirical data to demonstrate that increased standardization of curricula and assessments will not further isolate and ignore the needs of the learner. Absent this evidence, existing evidence demonstrates how a nationalized model of education will never work for all students. At the same time, greater resources should be allocated to combat out-of-school variables shown to influence student learning.

At the very least a more balanced approach to education policy is warranted. Two strategies might create greater balance in educational policy:

1. At the state and national level, education should be separated from politics. Appointments should not coincide with elections so as to protect these positions from political backlash. This strategy may allow for education polices to be created and implemented with purposeful patience.

2. Priority should be placed on education leaders being educators. Unlike other professions, the education profession has a preponderance of leaders that lack education experience at the school level. These leaders are more susceptible to the influence of ideologues and simple solutions to complex issues. They lack the experience necessary to understand the complexity of education issues at the school and district level and how issues impact diverse students and schools differently.

The stated goals of NCLB to raise the achievement for all students and close the various social and demographic achievement gaps are without question an admirable

goal. They are the essence of the American "way." Furthermore, the call for high standards can also be a great opportunity to promote greater clarity, quality, equity, and consistency of purpose in our educational endeavors at a national level. However, what kind of standards and assessments will incentivize educational practices in schools that value the individual needs of the learner and community equally with the needs of American society. Standards and assessments inspired by accountability only serve to promote conformity and prioritize national needs over the needs of individual learners and communities.

Essentialist policies are at odds with theories of learning. Essentialist policy-makers view progress in learning and student achievement as linear. Learning theories, however, demonstrate that learning is anything but linear. Learning is episodic, recycling, and cumulative (Sobol, 2003).

Policies of accountability imply society knows how to educate all students well and simply lacks the will. Interestingly, there is no precedent for any society raising all students to high levels of academic achievement regardless of social and demographic factors. Therefore, it is not clear that America, or any society, knows how to educate all students at increased levels. Essentialist policies promote "carrots" and "sticks" strategies, rationalizing that greater scrutiny will improve teaching and learning for all students. There are two glaring flaws in this approach to educational policy. First, research about organizational and human behavior convincingly contradicts the Theory X view of the worker. McGregor's Theory X and Theory Y research provides more evidence about the importance of leadership allowing for and expecting personal growth from workers in an organization versus a top-down, highly controlling leadership style.

McGregor's research-demonstrated managers' assumptions about people became self-fulfilling prophecies. McGregor found that managers (Theory Y) who see their workers as trustworthy and capable are significantly more effective than managers (Theory X) who see their workers as lazy and in need of control (Bolman & Deal 2008).

Essentialist theory is rooted in fear, the fear of "falling behind" and the desire to be superior. Essentialist theory misconstrues American values by misinterpreting the roots of Americanism. America's uniqueness and competitive advantage in the world springs from a long-standing commitment to individual protections and the limited role of government. Essentialist polices are more consistent with a nationalized education model and seek to inspire support by spreading fear about America becoming inferior. "Fear is a basic animal emotion experienced by all human beings. It has tremendous adaptive value, helping us to avoid threats and alerting us to danger and the need to react under stressful conditions – thereby helping us to survive (Marks, 1987)…Adolf Hitler is one of the most striking examples of the effective use of fear as a strategy to mobilize the public….But fear has been used to drive public policies in almost every society… Public policy and private lives have become fear-bound; fear has become the emotion through which public life is administered" (Zhao, 2009).

Progressive theories are based on the needs of the learner. They value the experience of students and view the learner in a holistic manner, valuing equally different academic disciplines. These theories require high levels of trust between stakeholders because these values cannot take hold within a climate where accountability and results are used as the means to the end. Rather, Progressive theories require the needs of the learner to determine the means to the end, which requires a much broader definition of

student achievement than standardized assessment data. Indeed, the Eight-Year Study demonstrated in convincing fashion how these Progressive theories best prepare students for college and career success.

The greatest disconnect between Essentialist policies and reality may be the widely accepted idea that the skills of innovative entrepreneurs will be required for success in the 21st Century much more than the skills of standardized laborers. This begs the question what type of curriculum designs should schools be implementing to nurture the development of innovation and entrepreneurial thinking in students. Furthermore, how do the Common Core Standards, a nationalized curriculum map, and subsequent high-stakes standardized assessments incentivize schools to implement a curriculum designed to teach the skills of innovative entrepreneurs? How are these policies consistent with the great tradition of American individualism? How do these policies empower impoverished communities and account for the influences of poverty on student learning? Or, will one unintended consequences of these policies be further oppression?

Today present day reforms like the Common Core Standards and Race to the Top are creating a new and more nationalized performance management system of American public education. "This new performance management system increases central control by top management freezing out the mediating functions of the middle layer of the organization. In the political realm this new approach to public administration increases the potential for a small group of centrally positioned elites to steer a whole system" (Mintrop & Sunderman, p. 353-354, 2009).

Worrisome is the lack of a research base for this nationalized approach to public

education. a policy which may lead to one of the largest social experiments ever conducted on children. Where is the evidence to suggest it will be successful?

Sputnik and Fear

By Sunday, October 6, 1957, most Americans had concluded that the beep, beep, beep from that thing called Sputnik, a manmade satellite that the Russians had supposedly sent into orbit on Friday, was not a hoax, not an electronic Potemkin Village, a product of what we would today call "special effects." Initially, the idea that Russian technology could surpass ours was unthinkable. And our brains' repression of Sputnik's reality was abetted by the implications that the putative orb carried: if they can send this thing over our heads, they can also attach an atomic bomb and drop it in our laps. "Soon they will be dropping bombs on us from space like kids dropping rocks from freeway overpasses," said Senate Majority Leader Lyndon Johnson. Writer Tom Wolfe described it this way: "Nothing less than *control of the heavens* was at stake. It was Armageddon, the final and decisive battle of the forces of good and evil." According to journalist Paul Dickson, ministers spoke of the Second Coming, and at least one said, "I wouldn't be surprised if He appeared today" (Bracey 2007).

Put simply, the fear-mongering had to stop. In fact, was America honestly "falling behind?" Was America "falling behind" the Soviet Union when Sputnik was launched in 1957? A review of documents from the Eisenhower Presidential Library suggests the United States made a strategic decision to allow the Soviet Union to launch its satellite first. In fact, the United States had the technology to launch a satellite into space in 1957. How then was the United States "falling behind" or inferior to the Soviet Union?

On the other hand, the Navy's Vanguard program was predicated on smaller rockets and lighter payloads, payloads that would clearly announce themselves as instruments of research. Alas, the Vanguard program was behind schedule and, ultimately, behind Sputnik. (Its post-Sputnik failures gave rise to headlines like "Kaputnik," "Flopnik," and "Dudnik"). Vanguard's schedule called for a November 1957 launch. This might well have caused the Russians to accelerate their own timetable. Vanguard finally pushed a 31-pound satellite into orbit on 31 March 1958. Eisenhower was casual about Sputnik. Indeed, his deputy secretary of defense, Donald Quarles, announced, "The Russians have, in fact, done us a 'good turn' unintentionally in establishing a doctrine of freedom of space." Eisenhower wrote, "We felt certain that we could get a great deal more information of all kinds out of the free use of space than they could." It was a wonderful doctrine that opened space up to exploration, but one for which educators paid a terrible price. For his part, Eisenhower was utterly perplexed that the success of Sputnik was seen to reflect a failed public school system" (Bracey, 2007, p. 121).

The United States had more than one rocket program in development in the 1950s and was working on many different technological initiatives within the military. President Eisenhower chose to be cautious about determining when to launch the United States' first rocket and satellite into space. It was a strategic decision based on foreign policy, not a result of inability to do so. President Eisenhower gave consideration to the impact of a space launch and its establishment of a doctrine that deep space was free and international space. Sputnik gave President Eisenhower the opportunity to define this doctrine of free space exploration, which he thought would benefit the United States.

The Sputnik event was hijacked by ideologues like Essentialist theorists Arthur Bestor, who argued loudly that Sputnik evidenced America was a less educated country. Bestor made these influential arguments despite facts like the average grade level of post World War I veterans was 6.8 compared to 10.5 for past World War II veterans. In fact, a mere twelve years after Sputnik, the United States landed men on the moon (Bestor, 2007).

How could the moon landing be possible if Americans were a less educated country? Since the end of World War II and the Cold War, America has become an unquestioned superpower. Again, how could this be possible if America suffered from an epidemic of failing public schools? If America's shortcomings are to be assigned to the failings of the public school systems, why are America's achievements not hailed as evidence for the success of the American public school system?

Overall Summary

Maylone (2002) found 56% of high school high-stakes standardized test data was explained by the percentage of lone-parent households, mean annual district household income, and the percentage of free- and reduced-lunch students in each high school community. This 2012 study, at the end of the original NCLB era, found the percentage of lone-parent households, percentage of households with at least a bachelor's degree, and percentage of economically disadvantaged families in a district explain 54% of 2009 NJ ASK Grade 3 Language Arts scores and 40% of 2009 NJ ASK Grade 3 Mathematics scores. Therefore, this study found no evidence to suggest the NCLB legislation positively impacted the achievement gap between students with greater resources and

those without. Still, Essentialist theories dominate the landscape of American public education policy with little evidence of their effectiveness.

Recommendations for Future Research

- Conduct a similar study using NJ ASK middle school data to determine if the theoretical framework from this study explains the greatest variance in scores
- Conduct a study of the school districts in which the greatest margins of error were identified by the predictive formula. Are school districts in which the actual score is greater than the predicted score implementing effective best practices, which may be causing the positive difference between predicted and actual scores? Similarly, are schools districts in which the actual score is less than the predicted score indeed guilty of underperforming and in need of correction action?
- Conduct a study to examine further the predictive power of the percentage of a population with a bachelor's degree in the household on student scores.
- Conduct this study at the school level to determine if the findings are corroborated or if different variables demonstrate greater influence than others at this level.
- Conduct a study to determine effective interventions for students influenced by the out-of-school variables found to explain significant variance in NJ ASK 3 scores.
- Conduct a study to determine effective early intervention programs for math learning.
- Conduct a study to determine effective strategies for improving student achievement in LAL while controlling for out-of-school variables identified in this study.

- Conduct a study to determine the influence of Common Core Standards on school level decisions about curriculum.
- Conduct a study of the assessment programs in progressive, private schools compared with assessment programs in the traditional public school system. How does each program impact curricula? How does each program impact the diverse needs of the student population? What happens to the graduates of these different programs?
- Since President Johnson's War on Poverty, billons of dollars have been spent on race-based initiatives with the intent of targeting issues of poverty. Conduct a study to determine how funding trends would shift if the focus of spending was tied more explicitly to issues of poverty.
- Conduct a study of schools where class rosters are balanced based on socio-economic factors compared with schools in which these factors are not considered. What differences are noted in student achievement levels?

Appendix A

Predictive District 2009 NJ ASK 3 LAL Scores from Theoretical Framework

Dist Code	DFG	District Name	% Passing LAL	Predicted Score Model A (% lone parent, % bach degree, % ED)	Margin of Error Model A
100	A	Asbury Park City	24.90	32.43	-7.53
110	A	Atlantic City	44.20	34.57	9.63
540	A	Bridgeton City	28.60	34.54	-5.94
590	A	Buena Regional	49.70	49.64	0.06
680	A	Camden City	19.90	26.45	-6.55
3880	A	City Of Orange Twp	51.00	45.36	5.64
950	A	Commercial Twp	50.80	40.27	10.53
1110	A	Dover Town	68.50	46.96	21.54
1210	A	East Orange	52.50	34.67	17.83
1300	A	Egg Harbor City	35.40	48.45	-13.05
1320	A	Elizabeth City	51.10	37.22	13.88
1460	A	Fairfield Twp	31.70	53.81	-22.11
1470	A	Fairview Boro	43.90	45.15	-1.25
2330	A	Irvington Township	41.30	34.16	7.14
2400	A	Keansburg Boro	45.10	36.20	8.90
2570	A	Lawrence Twp	57.10	52.49	4.61
3230	A	Millville City	41.40	47.26	-5.86
3570	A	Newark City	40.50	41.28	-0.78
3680	A	North Wildwood City	53.60	43.79	9.81
3970	A	Passaic City	29.10	47.75	-18.65
4010	A	Paterson City	33.00	34.01	-1.01
4020	A	Paulsboro Boro	28.70	37.73	-9.03
4070	A	Penns Grv-Carney's Pt Reg	35.80	39.96	-4.16
4090	A	Perth Amboy City	48.50	46.95	1.55
4180	A	Pleasantville City	41.20	36.64	4.56
4280	A	Quinton Twp	82.10	41.90	40.20
4630	A	Salem City	15.90	35.27	-19.37
4710	A	Seaside Heights Boro	64.50	47.36	17.14
5210	A	Trenton City	33.10	39.59	-6.49
5240	A	Union City	57.30	48.27	9.03

5390	A	Vineland City	48.10	48.29	-0.19
5670	A	West New York Town	54.50	49.30	5.20
5790	A	Wildwood City	43.60	42.44	1.16
5840	A	Woodbine Boro	23.50	34.69	-11.19
70	B	Alpha Boro	64.50	58.66	5.84
260	B	Bellmawr Boro	56.80	55.44	1.36
320	B	Berkeley Twp	74.00	52.95	21.05
490	B	Bound Brook Boro	40.80	52.29	-11.49
580	B	Brooklawn Boro	57.60	56.90	0.70
600	B	Burlington City	48.80	43.57	5.23
750	B	Carteret Boro	44.90	46.33	-1.43
880	B	Clementon Boro	59.20	38.25	20.95
890	B	Cliffside Park Boro	77.30	55.72	21.58
1020	B	Deerfield Twp	28.60	56.10	-27.50
1330	B	Elk Twp	71.10	58.28	12.82
1640	B	Freehold Boro	36.40	49.43	-13.03
1700	B	Garfield City	59.60	44.67	14.93
1730	B	Glassboro	43.70	49.45	-5.75
1770	B	Gloucester City	52.40	42.13	10.27
1850	B	Guttenberg Town	39.80	46.16	-6.36
1920	B	Haledon Boro	64.60	48.59	16.01
1960	B	Hammonton Town	44.20	53.81	-9.61
2060	B	Harrison Town	60.20	44.54	15.66
2390	B	Jersey City	40.60	51.70	-11.10
2410	B	Kearny Town	58.90	47.45	11.45
2500	B	Lakehurst Boro	50.00	52.43	-2.43
2660	B	Linden City	47.50	53.02	-5.52
2670	B	Lindenwold Boro	40.60	45.62	-5.02
2690	B	Little Egg Harbor Twp	52.90	54.93	-2.03
2740	B	Lodi Borough	52.30	55.23	-2.93
2770	B	Long Branch City	53.40	51.24	2.16
2940	B	Manchester Twp	49.40	55.53	-6.13
3050	B	Maurice River Twp	51.40	53.67	-2.27
3130	B	Middle Twp	54.70	54.92	-0.22
3300	B	Montague Twp	72.10	54.11	17.99
3350	B	Moonachie Boro	64.30	56.14	8.16
3430	B	Mount Holly Twp	42.50	47.26	-4.76
3480	B	Mullica Twp	52.20	56.94	-4.74
3490	B	National Park Boro	57.10	52.36	4.74
3540	B	New Hanover Twp	58.30	46.96	11.34
3610	B	North Bergen Twp	48.70	50.65	-1.95
3800	B	Ocean Gate Boro	40.90	53.69	-12.79
4050	B	Pemberton Twp	56.00	44.16	11.84
4100	B	Phillipsburg Town	58.30	37.22	21.08

4110	B	Pine Hill Boro	58.00	46.08	11.92
4160	B	Plainfield City	36.50	46.00	-9.50
4270	B	Prospect Park Boro	30.70	50.94	-20.24
4450	B	Riverside Twp	52.00	51.45	0.55
4590	B	Runnemede Boro	61.00	52.47	8.53
4850	B	South Bound Brook	59.60	55.14	4.46
5300	B	Upper Deerfield Twp	41.10	46.36	-5.26
5350	B	Ventnor City	60.50	52.93	7.57
5430	B	Wallington Boro	64.70	48.10	16.60
5740	B	Westville Boro	50.00	46.23	3.77
5760	B	Weymouth Twp	57.70	49.72	7.98
5800	B	Wildwood Crest Boro	77.80	52.00	25.80
5810	B	Winfield Twp	60.00	52.04	7.96
5860	B	Woodbury City	47.50	46.52	0.98
5900	B	Woodlynne Boro	29.20	44.11	-14.91
10	CD	Absecon City	54.30	52.38	1.92
185	CD	Barnegat Twp	68.60	55.27	13.33
220	CD	Bayonne City	63.10	50.76	12.34
250	CD	Belleville Town	51.80	52.00	-0.20
270	CD	Belmar Boro	69.60	54.63	14.97
340	CD	Berlin Twp	50.80	56.20	-5.40
500	CD	Bradley Beach Boro	44.40	50.24	-5.84
570	CD	Brigantine City	55.60	57.71	-2.11
860	CD	Clayton Boro	45.90	43.58	2.32
900	CD	Clifton City	49.50	54.03	-4.53
1030	CD	Delanco Twp	59.60	54.79	4.81
1080	CD	Dennis Twp	68.90	56.37	12.53
1100	CD	Deptford Twp	55.50	53.04	2.46
1230	CD	East Rutherford Boro	56.30	58.44	-2.14
1310	CD	Egg Harbor Twp	61.00	55.89	5.11
1345	CD	Elmwood Park	54.50	54.80	-0.30
1540	CD	Folsom Boro	34.10	58.20	-24.10
1570	CD	Franklin Boro	68.10	59.52	8.58
1590	CD	Franklin Twp	69.10	52.67	16.43
1690	CD	Galloway Twp	54.60	52.59	2.01
1860	CD	Hackensack City	70.40	49.13	21.27
1940	CD	Hamilton Twp	52.90	48.03	4.87
2190	CD	Hillside Twp	56.00	47.23	8.77
2270	CD	Hopewell Twp	52.20	52.21	-0.01
2430	CD	Keyport Boro	66.10	53.79	12.31
2710	CD	Little Ferry Boro	52.20	59.11	-6.91
2800	CD	Lower Alloways Creek	47.80	56.50	-8.70
2890	CD	Magnolia Boro	54.80	58.56	-3.76
2950	CD	Mannington Twp	78.30	54.84	23.46

3000	CD	Manville Boro	53.90	54.50	-0.60
3010	CD	Maple Shade Twp	48.70	48.71	-0.01
3280	CD	Monroe Twp	72.80	58.78	14.02
3420	CD	Mount Ephraim Boro	71.10	61.86	9.24
3500	CD	Neptune City	45.00	55.28	-10.28
3510	CD	Neptune Twp	42.50	51.26	-8.76
3590	CD	Newton Town	60.70	44.88	15.82
3650	CD	North Hanover Twp	57.80	58.05	-0.25
3770	CD	Oaklyn Boro	43.60	56.84	-13.24
3820	CD	Ocean Twp	54.10	61.32	-7.22
3860	CD	Oldmans Twp	70.00	49.66	20.34
3910	CD	Palisades Park	52.70	64.53	-11.83
4060	CD	Pennsauken Twp	50.40	46.60	3.80
4075	CD	Pennsville	48.30	48.67	-0.37
4150	CD	Pittsgrove Twp	51.20	57.91	-6.71
4290	CD	Rahway City	48.90	54.67	-5.77
4360	CD	Red Bank Boro	57.60	53.15	4.45
4790	CD	Somerdale Boro	69.60	58.27	11.33
4800	CD	Somers Point City	50.50	47.59	2.91
4830	CD	South Amboy City	50.00	51.36	-1.36
4870	CD	South Hackensack Twp	68.20	58.82	9.38
4920	CD	South River Boro	57.90	59.81	-1.91
5070	CD	Stow Creek Twp	55.00	58.89	-3.89
5200	CD	Totowa Boro	52.10	57.44	-5.34
5220	CD	Tuckerton Boro	71.00	48.16	22.84
5230	CD	Union Beach	76.40	53.28	23.12
5320	CD	Upper Pittsgrove Twp	78.90	56.07	22.83
5580	CD	Weehawken Twp	77.60	63.57	14.03
5820	CD	Winslow Twp	47.90	46.93	0.97
60	DE	Alloway Twp	62.70	58.93	3.77
150	DE	Audubon Boro	54.50	58.66	-4.16
280	DE	Belvidere Town	69.70	61.22	8.48
330	DE	Berlin Boro	81.20	58.69	22.51
410	DE	Bloomfield Twp	68.10	55.19	12.91
440	DE	Bogota Boro	47.40	57.46	-10.06
530	DE	Brick Twp	64.70	56.52	8.18
630	DE	Butler Boro	54.40	64.98	-10.58
740	DE	Carlstadt Boro	57.60	59.63	-2.03
1270	DE	Edgewater Boro	76.90	71.03	5.87
1280	DE	Edgewater Park Twp	65.60	50.29	15.31
1370	DE	Englewood City	55.00	50.75	4.25
1430	DE	Ewing Twp	49.70	56.97	-7.27
1520	DE	Florence Twp	61.50	51.76	9.74
1620	DE	Franklin Twp	48.80	59.63	-10.83

1710	DE	Garwood Boro	70.00	65.83	4.17
1780	DE	Gloucester Twp	61.50	53.14	8.36
1870	DE	Hackettstown	65.10	56.86	8.24
1930	DE	Hamburg Boro	70.30	59.71	10.59
2040	DE	Harmony Twp	79.10	62.01	17.09
2100	DE	Hawthorne Boro	70.30	62.28	8.02
2105	DE	Hazlet Twp	66.30	58.47	7.83
2370	DE	Jamesburg Boro	54.20	52.61	1.59
2420	DE	Kenilworth Boro	55.00	51.54	3.46
2480	DE	Lacey Twp	66.10	61.08	5.02
2540	DE	Laurel Springs Boro	86.40	52.62	33.78
2790	DE	Lopatcong Twp	80.00	54.00	26.00
2860	DE	Lyndhurst Twp	61.20	58.81	2.39
3020	DE	Margate City	80.00	59.71	20.29
3110	DE	Merchantville Boro	60.00	62.73	-2.73
3520	DE	Netcong Boro	64.30	50.18	14.12
3600	DE	North Arlington Boro	67.60	59.13	8.47
3670	DE	North Plainfield Boro	43.20	55.47	-12.27
3720	DE	Northfield City	69.20	56.12	13.08
3780	DE	Ocean City	68.90	51.95	16.95
3890	DE	Oxford Twp	61.30	51.11	10.19
3920	DE	Palmyra Boro	64.10	52.07	12.03
4190	DE	Plumsted Twp	62.90	53.18	9.72
4200	DE	Pohatcong Twp	60.60	58.47	2.13
4370	DE	Ridgefield Boro	62.90	53.00	9.90
4380	DE	Ridgefield Park Twp	85.70	57.21	28.49
4550	DE	Roselle Park Boro	73.40	57.31	16.09
4610	DE	Saddle Brook Twp	53.10	60.22	-7.12
4660	DE	Sayreville Boro	68.60	51.61	16.99
4730	DE	Secaucus Town	73.90	50.32	23.58
4930	DE	Southampton Twp	75.70	57.76	17.94
4970	DE	Spotswood Boro	54.10	58.92	-4.82
5020	DE	Stafford Twp	72.80	61.30	11.50
5080	DE	Stratford Boro	53.80	58.47	-4.67
5190	DE	Toms River Regional	70.20	55.71	14.49
5290	DE	Union Twp	59.70	54.45	5.25
5440	DE	Wanaque Boro	54.90	61.12	-6.22
5480	DE	Washington Boro	41.90	56.38	-14.48
5560	DE	Waterford Twp	55.20	59.52	-4.32
5620	DE	West Deptford Twp	61.90	55.89	6.01
5690	DE	West Patterson Boro	46.30	58.38	-12.08
5770	DE	Wharton Boro	49.30	56.43	-7.13
5780	DE	White Twp	73.10	59.23	13.87
5805	DE	Willingboro Twp	49.10	49.37	-0.27

5890	DE	Woodland Twp	41.70	59.79	-18.09
90	FG	Andover Reg	78.60	63.94	14.66
190	FG	Barrington Boro	56.90	53.06	3.84
300	FG	Bergenfield Boro	75.30	62.69	12.61
400	FG	Blairstown Twp	61.80	63.89	-2.09
420	FG	Bloomingdale Boro	45.20	62.68	-17.48
450	FG	Boonton Town	70.50	57.74	12.76
475	FG	Bordentown Regional	79.20	59.43	19.77
620	FG	Burlington Twp	52.80	59.33	-6.53
840	FG	Cinnaminson Twp	71.70	64.63	7.07
850	FG	Clark Twp	73.70	63.12	10.58
940	FG	Collingswood Boro	72.40	53.61	18.79
1060	FG	DELRAN	66.80	63.61	3.19
1130	FG	Dumont Boro	68.90	61.40	7.50
1140	FG	Dunellen Boro	71.60	52.15	19.45
1180	FG	East Greenwich Twp	69.50	63.56	5.94
1250	FG	Eastampton Twp	52.60	61.79	-9.19
1260	FG	Eatontown Boro	59.40	53.78	5.62
1550	FG	Fort Lee Boro	67.70	67.10	0.60
1560	FG	Frankford Twp	53.30	64.41	-11.11
1890	FG	Haddon Twp	73.00	63.59	9.41
1910	FG	Hainesport Twp	61.30	57.57	3.73
1950	FG	Hamilton Twp	63.70	57.76	5.94
2030	FG	Hardyston Twp	72.30	56.95	15.35
2080	FG	Hasbrouck Heights Boro	78.20	65.65	12.55
2210	FG	Hoboken City	45.70	59.31	-13.61
2220	FG	Holland Twp	76.90	63.99	12.91
2240	FG	Hopatcong	66.70	59.94	6.76
2250	FG	Hope Twp	52.00	59.14	-7.14
2290	FG	Howell Twp	72.10	63.62	8.48
2450	FG	Kingwood Twp	80.60	61.36	19.24
2650	FG	Lincoln Park Boro	77.50	62.07	15.43
2700	FG	Little Falls Twp	67.50	60.08	7.42
2750	FG	Logan Twp	65.50	54.38	11.12
2760	FG	Long Beach Island	61.50	62.48	-0.98
2850	FG	Lumberton Twp	77.20	56.97	20.23
2990	FG	Mantua Twp	72.40	61.05	11.35
3040	FG	Matawan-Aberdeen Regional	70.00	62.22	7.78
3060	FG	Maywood Boro	68.80	63.26	5.54
3140	FG	Middlesex Boro	61.80	55.85	5.95
3220	FG	Milltown Boro	70.30	66.71	3.59
3240	FG	Mine Hill Twp	68.00	55.88	12.12
3290	FG	Monroe Twp	70.60	63.50	7.10

3620	FG	North Brunswick Twp	70.70	54.12	16.58
3640	FG	North Haledon Boro	67.70	65.37	2.33
3730	FG	Northvale Boro	64.60	65.07	-0.47
3750	FG	Nutley Town	69.40	66.10	3.30
3840	FG	Ogdensburg Boro	63.60	60.73	2.87
3845	FG	Old Bridge Twp	69.60	63.77	5.83
4140	FG	Pitman Boro	59.30	57.69	1.61
4220	FG	Point Pleasant Beach Boro	67.30	62.04	5.26
4210	FG	Point Pleasant Boro	81.20	58.95	22.25
4230	FG	Pompton Lakes Boro	64.00	62.68	1.32
4440	FG	Riverdale Boro	79.30	63.96	15.34
4470	FG	Rochelle Park Twp	74.40	50.37	24.03
4480	FG	Rockaway Boro	65.10	62.51	2.59
4650	FG	Sandyston-Walpack Twp	84.20	61.70	22.50
4820	FG	Somerville Boro	50.00	57.50	-7.50
4880	FG	South Harrison Twp	78.60	61.66	16.94
4910	FG	South Plainfield Boro	68.00	59.94	8.06
4990	FG	Spring Lake Heights Boro	83.30	56.99	26.31
5040	FG	Stillwater Twp	70.90	58.86	12.04
5340	FG	Upper Twp	60.40	58.98	1.42
5360	FG	Vernon Twp	65.00	59.88	5.12
5500	FG	Washington Twp	70.90	57.20	13.70
5640	FG	West Long Branch Boro	63.20	59.97	3.23
5650	FG	West Milford Twp	70.60	64.85	5.75
5830	FG	Wood-Ridge Boro	68.80	66.37	2.43
5870	FG	Woodbury Heights Boro	56.50	53.68	2.82
20	GH	Alexandria Twp	62.50	70.17	-7.67
130	GH	Atlantic Highlands Boro	87.20	66.33	20.87
560	GH	Brielle Boro	75.90	67.62	8.28
800	GH	Cherry Hill Twp	79.60	64.36	15.24
830	GH	Chesterfield Twp	65.50	62.20	3.30
1040	GH	Delaware Twp	65.50	65.46	0.04
1190	GH	East Hanover Twp	87.20	67.74	19.46
1245	GH	East Windsor Regional	57.80	59.49	-1.69
1290	GH	Edison Twp	78.70	64.08	14.62
1360	GH	Emerson Boro	82.30	67.68	14.62
1450	GH	Fair Lawn Boro	74.20	67.81	6.39
1610	GH	Franklin Twp	55.10	64.78	-9.68
1630	GH	Fredon Twp	55.30	67.65	-12.35
1660	GH	Freehold Twp	79.90	64.96	14.94
1785	GH	Great Meadows Regional	79.80	61.88	17.92

1810	GH	Green Brook Twp	88.20	69.77	18.43
1880	GH	Haddon Heights Boro	64.50	66.49	-1.99
1980	GH	Hampton Twp	69.00	61.37	7.63
2070	GH	Harrison Twp	67.10	66.05	1.05
2140	GH	High Bridge Boro	64.30	61.40	2.90
2150	GH	Highland Park Boro	75.50	64.41	11.09
2180	GH	Hillsdale Boro	85.90	69.51	16.39
2380	GH	Jefferson Twp	67.30	62.62	4.68
2490	GH	Lafayette Twp	96.40	61.35	35.05
2530	GH	Lambertville City	75.00	55.26	19.74
2580	GH	Lawrence Twp	66.20	57.62	8.58
2620	GH	Leonia Boro	73.30	65.78	7.52
2680	GH	Linwood City	85.30	57.28	28.02
2920	GH	Manalapan-Englishtown Reg	72.50	66.07	6.43
2930	GH	Manasquan Boro	82.50	65.44	17.06
3160	GH	Middletown Twp	74.10	62.13	11.97
3170	GH	Midland Park Boro	72.60	68.19	4.41
3385	GH	Morris School District	64.80	64.34	0.46
3410	GH	Mount Arlington Boro	61.50	61.40	0.10
3450	GH	Mount Olive Twp	69.90	66.21	3.69
3830	GH	Oceanport Boro	72.70	66.12	6.58
3930	GH	Paramus Boro	77.00	62.82	14.18
3950	GH	Parsippany-Troy Hills Twp	72.20	67.29	4.91
4080	GH	Pequannock Twp	87.40	64.58	22.82
4130	GH	Piscataway Twp	63.30	59.28	4.02
4400	GH	Ringwood Boro	74.10	67.14	6.96
4460	GH	Riverton	83.70	67.08	16.62
4560	GH	Roxbury Twp	66.60	59.07	7.53
4600	GH	Rutherford Boro	73.80	66.36	7.44
5000	GH	Springfield Twp	81.00	64.39	16.61
5030	GH	Stanhope Boro	66.00	58.93	7.07
5130	GH	Tabernacle Twp	74.00	63.76	10.24
5150	GH	Teaneck Twp	62.30	62.64	-0.34
5185	GH	Tinton Falls	71.30	64.87	6.43
5270	GH	Union Twp	77.80	65.37	12.43
5310	GH	Upper Freehold Regional	73.60	59.02	14.58
5410	GH	Waldwick Boro	83.80	64.10	19.70
5420	GH	Wall Twp	76.60	63.73	12.87
5530	GH	Washington Twp	63.30	63.32	-0.02
5570	GH	Wayne Twp	70.80	67.13	3.67
5600	GH	West Amwell Twp	67.60	64.01	3.59

5680	GH	West Orange Town	70.00	63.26	6.74
5720	GH	Westampton	62.90	54.38	8.52
5755	GH	Westwood Regional	64.20	67.54	-3.34
30	I	Allamuchy Twp	53.30	69.33	-16.03
40	I	Allendale Boro	88.80	71.96	16.84
240	I	Bedminster Twp	86.20	69.02	17.18
310	I	Berkeley Heights Twp	79.90	71.16	8.74
370	I	Bethlehem Twp	81.70	71.27	10.43
460	I	Boonton Twp	85.20	68.66	16.54
510	I	Branchburg Twp	71.80	70.37	1.43
555	I	Bridgewater-Raritan Reg	73.80	66.37	7.43
640	I	Byram Twp	70.40	68.56	1.84
660	I	Caldwell-West Caldwell	80.00	66.61	13.39
760	I	Cedar Grove Twp	82.10	69.76	12.34
910	I	Clinton Town	70.20	71.77	-1.57
920	I	Clinton Twp	84.20	69.30	14.90
930	I	Closter Boro	80.30	69.14	11.16
945	I	Colts Neck Twp	80.60	69.47	11.13
980	I	Cranford Twp	73.30	68.19	5.11
990	I	Cresskill Boro	84.90	70.68	14.22
1070	I	Demarest Boro	92.10	67.85	24.25
1090	I	Denville Twp	78.40	69.40	9.00
1160	I	East Amwell Twp	81.60	65.01	16.59
1170	I	East Brunswick Twp	80.10	66.70	13.40
1380	I	Englewood Cliffs Boro	75.00	71.80	3.20
1420	I	Evesham Twp	73.80	65.01	8.79
1440	I	Fair Haven Boro	85.50	73.24	12.26
1510	I	Flemington-Raritan Reg	74.40	66.47	7.93
1530	I	Florham Park Boro	78.30	68.07	10.23
1580	I	Franklin Lakes Boro	77.10	69.53	7.57
1750	I	Glen Ridge Boro	86.60	73.42	13.18
1800	I	Green Twp	72.10	63.71	8.39
1840	I	Greenwich Twp	72.90	49.94	22.96
2000	I	Hanover Twp	79.70	68.40	11.30
2050	I	Harrington Park Boro	81.20	74.29	6.91
2090	I	Haworth Boro	91.10	71.82	19.28
2170	I	Hillsborough Twp	78.60	66.57	12.03
2230	I	Holmdel Twp	79.30	69.83	9.47
2280	I	Hopewell Valley Regional	75.20	69.17	6.03
2460	I	Kinnelon Boro	80.00	73.62	6.38
2590	I	Lebanon Boro	42.90	62.38	-19.48
2730	I	Livingston Twp	82.00	69.60	12.40
2870	I	Madison Boro	86.40	66.75	19.65

2900	I	Mahwah Twp	82.20	69.60	12.60
3030	I	Marlboro Twp	81.40	69.58	11.82
3070	I	Medford Lakes Boro	75.50	64.39	11.11
3080	I	Medford Twp	84.00	70.59	13.41
3120	I	Metuchen Boro	86.20	66.21	19.99
3200	I	Millstone Twp	74.30	61.45	12.85
3250	I	Monmouth Beach Boro	87.50	66.68	20.82
3310	I	Montclair Town	78.30	63.29	15.01
3330	I	Montvale Boro	80.30	66.34	13.96
3340	I	Montville Twp	78.20	70.18	8.02
3360	I	Moorestown Twp	81.00	63.42	17.58
3380	I	Morris Plains Boro	82.10	68.54	13.56
3440	I	Mount Laurel Twp	74.60	59.35	15.25
3470	I	Mountainside Boro	81.90	68.94	12.96
3560	I	New Providence Boro	82.40	71.96	10.44
3740	I	Norwood Boro	86.30	72.27	14.03
3760	I	Oakland Boro	83.60	67.85	15.75
3850	I	Old Tappan Boro	77.30	69.45	7.85
3940	I	Park Ridge Boro	73.10	71.10	2.00
4255	I	Princeton Regional	82.70	62.47	20.23
4310	I	Ramsey Boro	93.10	67.84	25.26
4330	I	Randolph Twp	73.80	71.13	2.67
4350	I	Readington Twp	88.50	66.34	22.16
4410	I	River Edge Boro	80.40	68.82	11.58
4430	I	River Vale Twp	88.80	70.83	17.97
5510	I	Robbinsville Twp	76.00	67.88	8.12
4490	I	Rockaway Twp	73.50	67.70	5.80
4530	I	Roseland Boro	81.00	68.86	12.14
4690	I	Sea Girt Boro	95.20	66.11	29.09
4770	I	Shrewsbury Boro	84.70	66.12	18.58
4815	I	Somerset Hills Regional	82.80	64.21	18.59
4860	I	South Brunswick Twp	74.50	65.04	9.46
4960	I	Sparta Twp	74.50	66.58	7.92
4980	I	Spring Lake Boro	92.30	69.61	22.69
5160	I	Tenafly Boro	88.40	66.28	22.12
5370	I	Verona Boro	84.10	69.43	14.67
5400	I	Voorhees Twp	79.30	58.04	21.26
5470	I	Warren Twp	83.10	71.69	11.41
5520	I	Washington Twp	75.40	62.30	13.10
5590	I	Wenonah Boro	77.40	70.50	6.90
5730	I	Westfield Town	84.10	68.44	15.66
5920	I	Wyckoff Twp	86.70	72.95	13.75
350	J	Bernards Twp	87.00	71.10	15.90
820	J	Chester Twp	79.70	68.50	11.20

970	J	Cranbury Twp	88.90	69.51	19.39
1400	J	Essex Fells Boro	85.30	71.07	14.23
1760	J	Glen Rock Boro	83.70	73.58	10.12
1900	J	Haddonfield Boro	81.20	71.30	9.90
2010	J	Harding Township	73.00	69.63	3.37
2200	J	Ho Ho Kus Boro	85.30	74.96	10.34
2720	J	Little Silver Boro	86.50	71.42	15.08
3090	J	Mendham Boro	82.10	70.83	11.27
3100	J	Mendham Twp	83.70	69.39	14.31
3190	J	Millburn Twp	86.20	70.15	16.05
3320	J	Montgomery Twp	78.40	70.82	7.58
3460	J	Mountain Lakes Boro	80.70	74.13	6.57
3630	J	North Caldwell Boro	90.00	66.22	23.78
4390	J	Ridgewood Village	88.30	68.19	20.11
4570	J	Rumson Boro	87.40	66.59	20.81
4620	J	Saddle River Boro	87.90	71.52	16.38
785	J	Sch Dist Of The Chathams	88.10	74.50	13.60
5180	J	Tewksbury Twp	86.00	70.79	15.21
5330	J	Upper Saddle River Boro	83.40	66.96	16.44
5715	J	W Windsor-Plainsboro Reg	91.20	47.67	43.53
Standard Deviation of Margin of Error					**10.53**

Appendix B

Predictive District 2009 NJ ASK 3 Math Scores from Theoretical Framework

Dist Code	DFG	District Name	% Passing Math	Predicted Score Model A (% lone parent, % bach degree, % ED)	Margin of Error Model A
100	A	Asbury Park City	39.9	49.46	-9.56
110	A	Atlantic City	63.1	51.45	11.65
540	A	Bridgeton City	44.7	51.83	-7.13
590	A	Buena Regional	67.1	65.26	1.84
680	A	Camden City	32	44.29	-12.29
3880	A	City Of Orange Twp	64.9	60.66	4.24
950	A	Commercial Twp	72.7	56.94	15.76
1110	A	Dover Town	86.6	63.07	23.53
1210	A	East Orange	65.2	51.41	13.79
1300	A	Egg Harbor City	47.9	64.27	-16.37
1320	A	Elizabeth City	68.6	54.08	14.52
1460	A	Fairfield Twp	33.3	68.63	-35.33
1470	A	Fairview Boro	67.2	61.15	6.05
2330	A	Irvington Township	49.3	50.98	-1.68
2400	A	Keansburg Boro	57.4	53.21	4.19
2570	A	Lawrence Twp	73.2	67.63	5.57
3230	A	Millville City	49.1	62.56	-13.46
3570	A	Newark City	52.3	57.03	-4.73
3680	A	North Wildwood City	82.1	59.77	22.33
3970	A	Passaic City	51	63.04	-12.04
4010	A	Paterson City	50.8	51.24	-0.44
4020	A	Paulsboro Boro	43.6	54.74	-11.14
4070	A	Penns Grv-Carney's Pt Reg	56.6	56.45	0.15
4090	A	Perth Amboy City	70.6	62.30	8.30
4180	A	Pleasantville City	57.1	53.47	3.63
4280	A	Quinton Twp	82.1	58.08	24.02
4630	A	Salem City	38.3	51.82	-13.52
4710	A	Seaside Heights Boro	74.2	62.71	11.49
5210	A	Trenton City	43.3	55.49	-12.19

5240	A	Union City	79.9	63.35	16.55
5390	A	Vineland City	63.5	63.55	-0.05
5670	A	West New York Town	75.1	64.20	10.90
5790	A	Wildwood City	60	57.95	2.05
5840	A	Woodbine Boro	47.1	51.76	-4.66
70	B	Alpha Boro	74.2	72.98	1.22
260	B	Bellmawr Boro	69.5	70.25	-0.75
320	B	Berkeley Twp	85.1	68.11	16.99
490	B	Bound Brook Boro	63.3	67.67	-4.37
580	B	Brooklawn Boro	57.6	71.58	-13.98
600	B	Burlington City	52.4	59.48	-7.08
750	B	Carteret Boro	67.1	62.17	4.93
880	B	Clementon Boro	83.7	54.37	29.33
890	B	Cliffside Park Boro	79.6	70.01	9.59
1020	B	Deerfield Twp	42.9	70.71	-27.81
1330	B	Elk Twp	80	73.01	6.99
1640	B	Freehold Boro	55.7	64.97	-9.27
1700	B	Garfield City	72.8	60.57	12.23
1730	B	Glassboro	59.4	64.54	-5.14
1770	B	Gloucester City	68.5	58.60	9.90
1850	B	Guttenberg Town	56.3	61.50	-5.20
1920	B	Haledon Boro	61.2	63.81	-2.61
1960	B	Hammonton Town	64.1	68.80	-4.70
2060	B	Harrison Town	68	60.59	7.41
2390	B	Jersey City	61	65.92	-4.92
2410	B	Kearny Town	76.2	63.31	12.89
2500	B	Lakehurst Boro	70	67.04	2.96
2660	B	Linden City	57.9	67.84	-9.94
2670	B	Lindenwold Boro	57.3	61.31	-4.01
2690	B	Little Egg Harbor Twp	65.2	69.61	-4.41
2740	B	Lodi Borough	73.8	69.85	3.95
2770	B	Long Branch City	67.8	65.98	1.82
2940	B	Manchester Twp	65.7	70.23	-4.53
3050	B	Maurice River Twp	94.3	69.03	25.27
3130	B	Middle Twp	72	69.33	2.67
3300	B	Montague Twp	85.7	69.16	16.54
3350	B	Moonachie Boro	64.3	70.48	-6.18
3430	B	Mount Holly Twp	62.3	62.63	-0.33
3480	B	Mullica Twp	79.1	71.50	7.60
3490	B	National Park Boro	74.3	67.85	6.45
3540	B	New Hanover Twp	100	63.40	36.60
3610	B	North Bergen Twp	74.6	65.62	8.98
3800	B	Ocean Gate Boro	86.4	68.58	17.82
4050	B	Pemberton Twp	69.8	60.13	9.67

4100	B	Phillipsburg Town	72.1	54.02	18.08
4110	B	Pine Hill Boro	74.8	62.02	12.78
4160	B	Plainfield City	50.3	61.38	-11.08
4270	B	Prospect Park Boro	52.8	65.88	-13.08
4450	B	Riverside Twp	70.3	66.39	3.91
4590	B	Runnemede Boro	79.2	67.27	11.93
4850	B	South Bound Brook	63.8	69.78	-5.98
5300	B	Upper Deerfield Twp	53.7	61.93	-8.23
5350	B	Ventnor City	75.4	67.41	7.99
5430	B	Wallington Boro	84.3	63.56	20.74
5740	B	Westville Boro	78.6	61.93	16.67
5760	B	Weymouth Twp	76.9	64.86	12.04
5800	B	Wildwood Crest Boro	92.6	66.21	26.39
5810	B	Winfield Twp	80	67.07	12.93
5860	B	Woodbury City	53.3	62.03	-8.73
5900	B	Woodlynne Boro	35.4	60.25	-24.85
10	CD	Absecon City	68.1	67.16	0.94
185	CD	Barnegat Twp	81	69.99	11.01
220	CD	Bayonne City	81.9	65.89	16.01
250	CD	Belleville Town	65.1	66.81	-1.71
270	CD	Belmar Boro	75	69.25	5.75
340	CD	Berlin Twp	76.9	71.07	5.83
500	CD	Bradley Beach Boro	82.1	64.81	17.29
570	CD	Brigantine City	73.6	71.98	1.62
860	CD	Clayton Boro	60.2	59.54	0.66
900	CD	Clifton City	58.9	68.73	-9.83
1030	CD	Delanco Twp	63.8	69.47	-5.67
1080	CD	Dennis Twp	90	70.54	19.46
1100	CD	Deptford Twp	68.2	67.94	0.26
1230	CD	East Rutherford Boro	70.8	72.54	-1.74
1310	CD	Egg Harbor Twp	75	70.54	4.46
1345	CD	Elmwood Park	74.7	69.59	5.11
1540	CD	Folsom Boro	61	72.56	-11.56
1570	CD	Franklin Boro	78.7	73.88	4.82
1590	CD	Franklin Twp	79.4	67.64	11.76
1690	CD	Galloway Twp	68.9	67.43	1.47
1860	CD	Hackensack City	72.6	64.23	8.37
1940	CD	Hamilton Twp	72.4	63.35	9.05
2190	CD	Hillside Twp	65.3	62.65	2.65
2270	CD	Hopewell Twp	59.7	67.17	-7.47
2430	CD	Keyport Boro	71	68.58	2.42
2710	CD	Little Ferry Boro	64.1	73.14	-9.04
2800	CD	Lower Alloways Creek	73.9	71.31	2.59
2890	CD	Magnolia Boro	59.5	72.85	-13.35

2950	CD	Mannington Twp	82.6	69.62	12.98
3000	CD	Manville Boro	70.6	69.45	1.15
3010	CD	Maple Shade Twp	61.2	64.04	-2.84
3280	CD	Monroe Twp	84.8	73.05	11.75
3420	CD	Mount Ephraim Boro	80	76.13	3.87
3500	CD	Neptune City	55	70.06	-15.06
3510	CD	Neptune Twp	56.5	66.14	-9.64
3590	CD	Newton Town	72.6	60.91	11.69
3650	CD	North Hanover Twp	69.8	72.66	-2.86
3770	CD	Oaklyn Boro	59	71.05	-12.05
3820	CD	Ocean Twp	78.7	74.96	3.74
3860	CD	Oldmans Twp	70	64.83	5.17
3910	CD	Palisades Park	78.3	77.53	0.77
4060	CD	Pennsauken Twp	69.1	62.46	6.64
4075	CD	Pennsville	65.5	64.41	1.09
4150	CD	Pittsgrove Twp	59.8	72.47	-12.67
4290	CD	Rahway City	62.9	69.22	-6.32
4360	CD	Red Bank Boro	74.1	67.27	6.83
4790	CD	Somerdale Boro	71.7	72.46	-0.76
4800	CD	Somers Point City	57.8	62.52	-4.72
4830	CD	South Amboy City	81.5	66.37	15.13
4870	CD	South Hackensack Twp	72.7	73.23	-0.53
4920	CD	South River Boro	72.1	74.13	-2.03
5070	CD	Stow Creek Twp	50	73.24	-23.24
5200	CD	Totowa Boro	69.9	71.83	-1.93
5220	CD	Tuckerton Boro	96.8	62.97	33.83
5230	CD	Union Beach	87.5	68.51	18.99
5320	CD	Upper Pittsgrove Twp	78.9	70.60	8.30
5580	CD	Weehawken Twp	88.1	76.36	11.74
5820	CD	Winslow Twp	67.9	62.79	5.11
60	DE	Alloway Twp	78.4	73.24	5.16
150	DE	Audubon Boro	86.4	72.91	13.49
280	DE	Belvidere Town	72.7	75.05	-2.35
330	DE	Berlin Boro	96.5	72.91	23.59
410	DE	Bloomfield Twp	75.8	69.62	6.18
440	DE	Bogota Boro	69.2	71.54	-2.34
530	DE	Brick Twp	73.5	70.98	2.52
630	DE	Butler Boro	80.7	78.35	2.35
740	DE	Carlstadt Boro	66.1	73.97	-7.87
1270	DE	Edgewater Boro	95.5	83.22	12.28
1280	DE	Edgewater Park Twp	78.5	65.43	13.07
1370	DE	Englewood City	61.6	65.84	-4.24
1430	DE	Ewing Twp	56.3	71.30	-15.00
1520	DE	Florence Twp	70.1	66.64	3.46

1620	DE	Franklin Twp	58.5	73.56	-15.06
1710	DE	Garwood Boro	82.5	79.39	3.11
1780	DE	Gloucester Twp	73.5	67.99	5.51
1870	DE	Hackettstown	90.7	71.20	19.50
1930	DE	Hamburg Boro	64.9	73.73	-8.83
2040	DE	Harmony Twp	83.7	76.16	7.54
2100	DE	Hawthorne Boro	82.4	75.82	6.58
2105	DE	Hazlet Twp	72.1	72.78	-0.68
2370	DE	Jamesburg Boro	81.7	67.33	14.37
2420	DE	Kenilworth Boro	76.2	66.84	9.36
2480	DE	Lacey Twp	74.4	75.14	-0.74
2540	DE	Laurel Springs Boro	95.5	67.13	28.37
2790	DE	Lopatcong Twp	82.6	69.10	13.50
2860	DE	Lyndhurst Twp	68.9	73.01	-4.11
3020	DE	Margate City	98.2	73.44	24.76
3110	DE	Merchantville Boro	91.4	75.79	15.61
3520	DE	Netcong Boro	71.4	65.66	5.74
3600	DE	North Arlington Boro	86	73.55	12.45
3670	DE	North Plainfield Boro	55.1	70.10	-15.00
3720	DE	Northfield City	83.8	70.73	13.07
3780	DE	Ocean City	79.7	66.45	13.25
3890	DE	Oxford Twp	77.4	66.73	10.67
3920	DE	Palmyra Boro	89.1	66.67	22.43
4190	DE	Plumsted Twp	86	68.45	17.55
4200	DE	Pohatcong Twp	57.6	72.93	-15.33
4370	DE	Ridgefield Boro	79.1	67.81	11.29
4380	DE	Ridgefield Park Twp	93.5	71.27	22.23
4550	DE	Roselle Park Boro	79	71.47	7.53
4610	DE	Saddle Brook Twp	77	74.31	2.69
4660	DE	Sayreville Boro	85.5	66.89	18.61
4730	DE	Secaucus Town	87	65.57	21.43
4930	DE	Southampton Twp	84.3	72.47	11.83
4970	DE	Spotswood Boro	72.1	73.32	-1.22
5020	DE	Stafford Twp	85.4	75.28	10.12
5080	DE	Stratford Boro	71.8	72.39	-0.59
5190	DE	Toms River Regional	86.2	70.35	15.85
5290	DE	Union Twp	72.4	69.13	3.27
5440	DE	Wanaque Boro	77.5	75.07	2.43
5480	DE	Washington Boro	72.4	70.76	1.64
5560	DE	Waterford Twp	67.6	73.91	-6.31
5620	DE	West Deptford Twp	75.5	70.21	5.29
5690	DE	West Patterson Boro	68.6	72.47	-3.87
5770	DE	Wharton Boro	86.3	70.55	15.75
5780	DE	White Twp	65.4	73.92	-8.52

5805	DE	Willingboro Twp	59.8	64.62	-4.82
5890	DE	Woodland Twp	50	74.12	-24.12
90	FG	Andover Reg	92.9	77.54	15.36
190	FG	Barrington Boro	81.5	67.97	13.53
300	FG	Bergenfield Boro	89.1	76.29	12.81
400	FG	Blairstown Twp	69.6	77.44	-7.84
420	FG	Bloomingdale Boro	78.1	76.81	1.29
450	FG	Boonton Town	78.7	71.62	7.08
475	FG	Bordentown Regional	85.2	73.30	11.90
620	FG	Burlington Twp	76.4	73.33	3.07
840	FG	Cinnaminson Twp	82.8	78.03	4.77
850	FG	Clark Twp	79.6	76.74	2.86
940	FG	Collingswood Boro	74.2	67.92	6.28
1060	FG	DELRAN	65.4	77.16	-11.76
1130	FG	Dumont Boro	79.8	74.95	4.85
1140	FG	Dunellen Boro	83	67.07	15.93
1180	FG	East Greenwich Twp	78.9	77.12	1.78
1250	FG	Eastampton Twp	60.5	75.64	-15.14
1260	FG	Eatontown Boro	73.5	68.35	5.15
1550	FG	Fort Lee Boro	84.4	79.91	4.49
1560	FG	Frankford Twp	72.1	77.85	-5.75
1890	FG	Haddon Twp	87.6	77.04	10.56
1910	FG	Hainesport Twp	86.7	71.62	15.08
1950	FG	Hamilton Twp	65.8	72.13	-6.33
2030	FG	Hardyston Twp	78.3	71.29	7.01
2080	FG	Hasbrouck Heights Boro	89.1	78.96	10.14
2210	FG	Hoboken City	63	72.39	-9.39
2220	FG	Holland Twp	63.1	77.35	-14.25
2240	FG	Hopatcong	77.4	74.14	3.26
2250	FG	Hope Twp	76	73.07	2.93
2290	FG	Howell Twp	87.1	77.19	9.91
2450	FG	Kingwood Twp	88.9	75.16	13.74
2650	FG	Lincoln Park Boro	92.5	75.64	16.86
2700	FG	Little Falls Twp	73.5	74.11	-0.61
2750	FG	Logan Twp	72.9	69.11	3.79
2760	FG	Long Beach Island	69.2	75.94	-6.74
2850	FG	Lumberton Twp	85.3	71.55	13.75
2990	FG	Mantua Twp	88.1	74.89	13.21
3040	FG	Matawan-Aberdeen Regional	87.2	75.84	11.36
3060	FG	Maywood Boro	72	76.70	-4.70
3140	FG	Middlesex Boro	86.7	70.52	16.18
3220	FG	Milltown Boro	82.4	79.84	2.56
3240	FG	Mine Hill Twp	78	70.62	7.38

3290	FG	Monroe Twp	86.3	77.29	9.01
3620	FG	North Brunswick Twp	80	68.77	11.23
3640	FG	North Haledon Boro	59.8	78.45	-18.65
3730	FG	Northvale Boro	72.7	78.66	-5.96
3750	FG	Nutley Town	89.2	79.20	10.00
3840	FG	Ogdensburg Boro	100	74.89	25.11
3845	FG	Old Bridge Twp	80.6	77.20	3.40
4140	FG	Pitman Boro	86.4	71.96	14.44
4220	FG	Point Pleasant Beach Boro	60	75.72	-15.72
4210	FG	Point Pleasant Boro	91.6	72.52	19.08
4230	FG	Pompton Lakes Boro	76.3	76.18	0.12
4440	FG	Riverdale Boro	86.2	77.76	8.44
4470	FG	Rochelle Park Twp	81.4	64.77	16.63
4480	FG	Rockaway Boro	76.2	76.08	0.12
4650	FG	Sandyston-Walpack Twp	94.7	75.54	19.16
4820	FG	Somerville Boro	62.7	71.99	-9.29
4880	FG	South Harrison Twp	92.9	75.58	17.32
4910	FG	South Plainfield Boro	76.8	74.03	2.77
4990	FG	Spring Lake Heights Boro	77.8	70.97	6.83
5040	FG	Stillwater Twp	85.5	72.77	12.73
5340	FG	Upper Twp	75.5	73.13	2.37
5360	FG	Vernon Twp	69.9	73.92	-4.02
5500	FG	Washington Twp	86.8	72.50	14.30
5640	FG	West Long Branch Boro	82.4	74.11	8.29
5650	FG	West Milford Twp	86.1	78.35	7.75
5830	FG	Wood-Ridge Boro	58.5	79.64	-21.14
5870	FG	Woodbury Heights Boro	82.6	68.43	14.17
20	GH	Alexandria Twp	72.2	83.02	-10.82
130	GH	Atlantic Highlands Boro	84.6	79.30	5.30
560	GH	Brielle Boro	94	80.33	13.67
800	GH	Cherry Hill Twp	92.5	77.70	14.80
830	GH	Chesterfield Twp	75.9	75.96	-0.06
1040	GH	Delaware Twp	64.3	78.69	-14.39
1190	GH	East Hanover Twp	87.2	80.94	6.26
1245	GH	East Windsor Regional	76.2	73.48	2.72
1290	GH	Edison Twp	87.1	77.48	9.62
1360	GH	Emerson Boro	93.8	80.68	13.12
1450	GH	Fair Lawn Boro	87.7	80.69	7.01
1610	GH	Franklin Twp	71.8	78.44	-6.64
1630	GH	Fredon Twp	71.1	80.62	-9.52
1660	GH	Freehold Twp	85.3	78.31	6.99
1785	GH	Great Meadows	90.3	75.61	14.69

		Regional			
1810	GH	Green Brook Twp	89.3	82.61	6.69
1880	GH	Haddon Heights Boro	82.3	79.27	3.03
1980	GH	Hampton Twp	84.5	75.41	9.09
2070	GH	Harrison Twp	77.5	79.13	-1.63
2140	GH	High Bridge Boro	78.6	74.83	3.77
2150	GH	Highland Park Boro	85.7	77.87	7.83
2180	GH	Hillsdale Boro	87.2	82.29	4.91
2380	GH	Jefferson Twp	84.8	76.27	8.53
2490	GH	Lafayette Twp	100	75.38	24.62
2530	GH	Lambertville City	95.8	69.75	26.05
2580	GH	Lawrence Twp	80	71.83	8.17
2620	GH	Leonia Boro	80.4	78.50	1.90
2680	GH	Linwood City	94	71.74	22.26
2920	GH	Manalapan-Englishtown Reg	86	79.31	6.69
2930	GH	Manasquan Boro	85	78.72	6.28
3160	GH	Middletown Twp	87	75.87	11.13
3170	GH	Midland Park Boro	84.9	81.15	3.75
3385	GH	Morris School District	75	77.43	-2.43
3410	GH	Mount Arlington Boro	74.4	75.02	-0.62
3450	GH	Mount Olive Twp	83.5	79.18	4.32
3830	GH	Oceanport Boro	81.8	79.15	2.65
3930	GH	Paramus Boro	89.1	76.47	12.63
3950	GH	Parsippany-Troy Hills Twp	84.9	80.19	4.71
4080	GH	Pequannock Twp	96.6	77.94	18.66
4130	GH	Piscataway Twp	84.8	73.26	11.54
4400	GH	Ringwood Boro	91.9	80.28	11.62
4460	GH	Riverton	86	80.30	5.70
4560	GH	Roxbury Twp	79.8	73.42	6.38
4600	GH	Rutherford Boro	80.2	79.28	0.92
5000	GH	Springfield Twp	87.2	77.98	9.22
5030	GH	Stanhope Boro	76.6	73.20	3.40
5130	GH	Tabernacle Twp	67.5	77.61	-10.11
5150	GH	Teaneck Twp	82.1	75.99	6.11
5185	GH	Tinton Falls	80.5	78.10	2.40
5270	GH	Union Twp	90.7	79.01	11.69
5310	GH	Upper Freehold Regional	86	73.37	12.63
5410	GH	Waldwick Boro	92.3	77.48	14.82
5420	GH	Wall Twp	91	76.98	14.02
5530	GH	Washington Twp	72.2	76.82	-4.62
5570	GH	Wayne Twp	83.1	80.14	2.96

5600	GH	West Amwell Twp	78.4	77.64	0.76
5680	GH	West Orange Town	74	76.61	-2.61
5720	GH	Westampton	71.1	69.02	2.08
5755	GH	Westwood Regional	78.6	80.47	-1.87
30	I	Allamuchy Twp	71.1	81.89	-10.79
40	I	Allendale Boro	95.5	84.25	11.25
240	I	Bedminster Twp	90.9	81.62	9.28
310	I	Berkeley Heights Twp	84.2	83.51	0.69
370	I	Bethlehem Twp	92.7	83.67	9.03
460	I	Boonton Twp	86.9	81.36	5.54
510	I	Branchburg Twp	83.6	83.02	0.58
555	I	Bridgewater-Raritan Reg	88.3	79.53	8.77
640	I	Byram Twp	88	81.43	6.57
660	I	Caldwell-West Caldwell	81.5	79.39	2.11
760	I	Cedar Grove Twp	79.3	82.21	-2.91
910	I	Clinton Town	72.3	83.81	-11.51
920	I	Clinton Twp	94.7	81.98	12.72
930	I	Closter Boro	91.8	81.73	10.07
945	I	Colts Neck Twp	65.2	82.02	-16.82
980	I	Cranford Twp	84.4	81.00	3.40
990	I	Cresskill Boro	91.4	83.10	8.30
1070	I	Demarest Boro	87.7	80.50	7.20
1090	I	Denville Twp	84.8	81.95	2.85
1160	I	East Amwell Twp	95.9	78.42	17.48
1170	I	East Brunswick Twp	89	79.66	9.34
1380	I	Englewood Cliffs Boro	86.7	83.76	2.94
1420	I	Evesham Twp	80	78.07	1.93
1440	I	Fair Haven Boro	83.9	85.16	-1.26
1510	I	Flemington-Raritan Reg	87.9	79.46	8.44
1530	I	Florham Park Boro	89.6	80.84	8.76
1580	I	Franklin Lakes Boro	89.2	81.90	7.30
1750	I	Glen Ridge Boro	89	85.29	3.71
1800	I	Green Twp	89.7	77.02	12.68
1840	I	Greenwich Twp	81.3	64.73	16.57
2000	I	Hanover Twp	87.5	81.20	6.30
2050	I	Harrington Park Boro	85.5	86.27	-0.77
2090	I	Haworth Boro	98.2	84.25	13.95
2170	I	Hillsborough Twp	89.1	79.43	9.67
2230	I	Holmdel Twp	83.6	82.58	1.02
2280	I	Hopewell Valley Regional	87.1	81.88	5.22
2460	I	Kinnelon Boro	83.5	85.64	-2.14
2590	I	Lebanon Boro	78.6	75.71	2.89
2730	I	Livingston Twp	90.2	82.10	8.10

2870	I	Madison Boro	89.9	79.69	10.21
2900	I	Mahwah Twp	89.7	82.01	7.69
3030	I	Marlboro Twp	90.3	82.29	8.01
3070	I	Medford Lakes Boro	94.3	76.99	17.31
3080	I	Medford Twp	91.1	82.99	8.11
3120	I	Metuchen Boro	92.2	79.12	13.08
3200	I	Millstone Twp	82.7	74.98	7.72
3250	I	Monmouth Beach Boro	81.2	79.65	1.55
3310	I	Montclair Town	85.1	76.26	8.84
3330	I	Montvale Boro	94	79.38	14.62
3340	I	Montville Twp	90.6	82.66	7.94
3360	I	Moorestown Twp	88.7	76.91	11.79
3380	I	Morris Plains Boro	92.9	81.43	11.47
3440	I	Mount Laurel Twp	86.2	73.31	12.89
3470	I	Mountainside Boro	94	81.89	12.11
3560	I	New Providence Boro	91.6	84.02	7.58
3740	I	Norwood Boro	88	84.60	3.40
3760	I	Oakland Boro	85.7	80.64	5.06
3850	I	Old Tappan Boro	81	82.18	-1.18
3940	I	Park Ridge Boro	84.3	83.41	0.89
4255	I	Princeton Regional	89.7	76.17	13.53
4310	I	Ramsey Boro	96.1	80.74	15.36
4330	I	Randolph Twp	84.7	83.49	1.21
4350	I	Readington Twp	95.1	79.37	15.73
4410	I	River Edge Boro	89	81.40	7.60
4430	I	River Vale Twp	90.1	83.36	6.74
5510	I	Robbinsville Twp	86.6	80.67	5.93
4490	I	Rockaway Twp	83.4	80.62	2.78
4530	I	Roseland Boro	88.9	81.34	7.56
4690	I	Sea Girt Boro	95.2	78.77	16.43
4770	I	Shrewsbury Boro	91.5	79.26	12.24
4815	I	Somerset Hills Regional	88.3	77.51	10.79
4860	I	South Brunswick Twp	86.8	78.34	8.46
4960	I	Sparta Twp	89.3	79.78	9.52
4980	I	Spring Lake Boro	92.3	82.54	9.76
5160	I	Tenafly Boro	93.8	79.23	14.57
5370	I	Verona Boro	84.2	82.11	2.09
5400	I	Voorhees Twp	87.6	72.12	15.48
5470	I	Warren Twp	89.2	84.03	5.17
5520	I	Washington Twp	90.4	75.84	14.56
5590	I	Wenonah Boro	87.1	82.67	4.43
5730	I	Westfield Town	93.6	81.26	12.34
5920	I	Wyckoff Twp	92.6	84.96	7.64
350	J	Bernards Twp	91.5	83.42	8.08

820	J	Chester Twp	95.6	81.11	14.49
970	J	Cranbury Twp	88.9	81.98	6.92
1400	J	Essex Fells Boro	94.1	83.21	10.89
1760	J	Glen Rock Boro	91.3	85.71	5.59
1900	J	Haddonfield Boro	90	83.54	6.46
2010	J	Harding Township	73	81.91	-8.91
2200	J	Ho Ho Kus Boro	93.3	86.72	6.58
2720	J	Little Silver Boro	90.6	83.39	7.21
3090	J	Mendham Boro	95.5	83.25	12.25
3100	J	Mendham Twp	90.4	81.89	8.51
3190	J	Millburn Twp	94.6	82.62	11.98
3320	J	Montgomery Twp	83.8	83.24	0.56
3460	J	Mountain Lakes Boro	87.8	85.93	1.87
3630	J	North Caldwell Boro	94.4	79.39	15.01
4390	J	Ridgewood Village	94.1	80.88	13.22
4570	J	Rumson Boro	88.3	79.46	8.84
4620	J	Saddle River Boro	97	83.79	13.21
785	J	Sch Dist Of The Chathams	96.5	86.37	10.13
5180	J	Tewksbury Twp	94.6	83.23	11.37
5330	J	Upper Saddle River Boro	86.3	79.87	6.43
5715	J	W Windsor-Plainsboro Reg	94	63.02	30.98
Standard Deviation					**10.07**

References

Adler, M. J. (1982). *The paideia proposal: An educational manifesto.* New York, NY: Simon & Schuster.

Aiken, W. M. (1942). *The story of the eight-year study.* New York, NY: Harper and Brothers.

Alspaugh, J. (1991). Out-of-school environmental factors and elementary school achievement in mathematics and reading. *Journal of Research and Development in Education, 24*(3), 53-55.

Amato, P. R., & Keith, B. (1991). Parental divorce and adult well-being: A meta-analysis. *Journal of Marriage and Family, 53*(1), 43-58.

American Psychological Association. (1999). *The standards for educational and psychological testing.* Washington, D.C.: American Psychological Association.

Amrein, A. L., & Berliner, C. D. (2002, March). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives, 10*(18). Retrieved from http://epaa.asu.edu/epaa/v10n18

Amrein, A. L., & Berliner, D. C. (2002, December). *An analysis of some unintended and negative consequence of high-stakes testing.* East Lansing, MI: The Great Lakes Center for Education Research & Practice.

Astone, N. M., & McLanahan, S. S. (1991). Family Structure, Parental Practices, and High School Completion. *American Sociological Assocition,* 56(3), 309-320.

Bagley, W. C. (1907). *The educative process.* London: The MacMillan Company.

Bagley, W. C. (1912). *Educational values*. New York: Mavmillan.

Bagley, W. C. (1919). Your education is my business. *Music Supervisors' Journal,* 6(2), 30.

Bagley, W. C. (1934). *Education and emergent man: A theory of education with particular application to public education in the United State*s. New York, NY: Thomas Nelson and Sons.

Bagley, W. C. (1938). An essentialist's platform for the advancement of American education. *Educational Administration and Supervision, 24*(April), 241.

Bagley, W. C. (1939). The significance of the essentialist movement in educational theory. *The Classical Journal, 3*4(6), 326.

Bennett, R. V. (1972). *Cardinal principles report: An educational classic*. College Park, Maryland: Department of Administration, Supervision and Curriculum, University of Maryland.

Berliner, D. C., & Nichols, S. L. (2005, March). *The inevitable corruption of indicators and educators through high-stakes testing*. East Lansing, MI: The Great Lakes Center for Education Research & Practice.

Bestor, A. (1955). *The restoration of learning: A program for redeeming the unfulfilled promise of American educatio*n. New York, NY: Knopf.

Blau, D. M. (1999). The effect of income on child development. *The Review of Economics and Statistics, 8*1(2), 261-276.

Bracey, Gerald W. The first time "everything changed": The 17th Bracey report on the condition of public education, Vol. 89, No. 02, October 2007, pp. 119-136.

Braun, H. (2004). Reconsidering the impact of high-stakes testing. *Education Policy Analysis Archives, 12*(1). Retrieved from http://epaa.asu.edu/epaa/v12n1

Bolman, L.G., & Deal, T.E. (2008) Reframing organizations: Artisry, choice, and leadership. San Francisco, CA: John Wiley & Sons.

Carlson, M. J., & Corcoran, M. E. (2001, March). Family structure and children's behavioral and cognitive outcomes. *Journal of Marriage and Family, 63*(3), 779.

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

Center on Education Policy. (2009, October). *Are achievement gaps closing and is achievement rising for all*? Washington, D.C.: Center on Education Policy.

Center on Education Policy. (2010, September). *Rising scores on state tests and NAEP.* Washington, D.C.: Center on Education Policy.

Center on Education Policy. (2010, December). *Slow and uneven progress in narrowing gaps*. Washington, D.C.: Center on Education Policy.

Center on Education Policy. (2011, April). *State test score trends through 2008-09, part 3: Student achievement at 8th grade*. Washington, D.C.: Center on Education Policy.

Chudowsky, N., & Chudowsky, V. (2010, September). *State test score trends through 2008-09, part 1*. Washington, D.C.: Center on Education Policy.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94-95.

Coleman, J. S. (1966). *Equality of educational opportunity*. Ann Arbor, MI: Inter-University Consortium for Policial and Social Research. 2007-04-27. doi:10.3886/ICPSRO6389.v3

Coleman, W. T., Selby, C.

C, et al. (1983). *Educating Americans for the 21st century: A plan of action for improving mathematics, science and technology education for all American elementary and secondary students so their achievement is the best in the world by 1995.* Washington, D.C.: National Science Foundation.

Conant, J. B. (1964). *Shaping educational policy*. New York, NY: McGraw-Hill.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.

Cuban, L. (1993). *How teachers taught: Consistency and change in American classrooms 1880-1990*. New York, NY: Teachers College Press.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology, 19*(2), 294.

Dawson, D. A. (1991). Family structure and children's health and well-being: Data from the 1988 national health interview survey on child health. *Journal of Marriage and Family, 53*(3), 573-584.

Dewey, J. (1897). *The significance of the problem of knowledge*. Chicago, IL: The University of Chicago Press.

Dewey, J. (1900). Psychology and social practice. *Psychological Review, 7*(2), 105-124.

Dewey, J. (1902). *The child and the curriculum*. Chicago, IL: The University of Chicago Press.

Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. New York, NY: MacMillan Company.

Dewey, J. (1938). *Experience and education*. New York, NY: Kappa Delta Pi.

Dorn, S. (1998). The political legacy of school accountability systems. *Education Policy Analysis Archives,* 6(1)

Downey, D. B. (1995). Understanding academic achievement among children in stephouseholds: The role of parental resources, sex of stepparent, and sex of child. *University of North Carolina Press,* 73(3), 875-894.

Elmore, R. F. (2004). *School reform from the inside out: Policy, practice, and performance*. Cambridge, MA: Harvard Education Press.

Fryer, R. G., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *The Review of Economics and Statistics, 86*(2), 447-464.

Gamorna, A., & Long, D. A. (2006). Equality of educational opportunity: A 40-year retrospective (WCER manuscript in preparation No. 2006-9). WI: Wisconsin Center for Educational Research,

Gay, L. R., Mills, G. E., & Airasian, P. W. (Eds.). (2009). *Educational research: Competencies for analysis and applications* (9th ed.). New Jersey: Pearson.

Gay, L.R., Mills, G. E., & Airasian, P.W. (2009). In 9th (Ed.), *Educational research: Competencies for analysis and applications*. Upper Saddle River, NJ: Prentice Hall.

Goals 2000: Educate America Act, 103-227 (1994).

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996a). The effect of school resources on student achievement. *Review of Educational Research, 66*, 361.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996b). Interpreting research on school resources and student achievement: A rejoinder to Hanushek. *Review of Educational Research, 66*, 441.

Guyatt, G. H., & Rennie, D. (2002). *User's guide to the medical literature*. Chicago, IL: AMA Press.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature, 24*, 1141.

Hanushek, E. A. (1996). A more complete picture of school resource policies. *Review of Educational Research, 66*, 367.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis, 19*, 141.

Hanushek, E. A., & Raymond, M. E. (2003). High-stakes research: The campaign against accountability has brought negative anecdotes and deeply flawed research. Solid analysis reveals a brighter picture. *Education Next,* 3(3)

Hanushek, E. A., & Raymond, M. E. (2004). The effect of school accountability systems on the level and distribution of student achievement. *Journal of the European Economic Association,* 2(2), 406-415.

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.

Hanushek, E. A., Raymond, M. E., & Rivkin, S. G. (2004). *Does it matter how we judge school quality*? Salt Lake City, Utah: Hoover Institution.

Hauser, R. M., & Sewell, W. H. (1986). Family effects in simple models of education, occupational status, and earnings: Findings from the Wisconsin and Kalamazoo studies. *Journal of Labor Economics,* 4(3), 83-115.

Hernandez, D. J. (2011). *How third-grade reading skills and poverty influence high school graduation*. NY: Hunter College..

Hinkle, D. E., Wiersman, W., & Jurs, S. G. (2003). Applied statistics for the behavioral sciences. *Peabody Journal of Education, 67*(3), 3.

Hirsch, E. D. (1987). *Cultural literacy: What every American needs to know*. New York, NY: Houghton Mifflin Company.

Hirsch, E. D. (1996). *The schools we need and why we don't have them*. New York, NY: Anchor Books.

Hull, C. L. (1943). *Principles of behavior*. New York, NY: Appleton-Century-Crofts.

Hull, C. L. (1951). *Essentials of behavior*. New Haven, CT: Yale University Press.

Irvin, L. K., Horner, R. H., Ingram, K., Todd, A., Sugai, G., Sampson, N. K., & Boland, J. (2006). Using office discipline referral data for decision making about student behavior in elementary and middle schools: An empirical evaluation of validity. *Journal of Positive Behavior Interventions,* 8(1), 10.

Jencks, C., & Phillips, M. (1998). The black-white test scope gap: Why it persists and what can be done. *The Brookings Review, 16*(2), 24-27.

Jencks, C., Smith, M., Acland, H., Bane, M.S., Cohen, D., Gintis, H., et al. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York, NY: Basic Books.

Jimerson, S., Egeland, B., & Teo, A. (1999). A longitudinal study of achievement trajectories factors associated with change. *Journal of Educational Psychology, 91*, 116.

Johanningmeier, E. V. (1969). William Chandler Bagley's changing views on the relationship between psychology and education. *History of Education Quarterly, 9*(1), 3.

Johnson, B. (2001). Toward a new classification of nonexperimental quantitative research. *Educational Researcher, 30*(3), 3-12.

Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education.* Washington, D.C.: Joint Committee on Testing Practices.

Jones, M. (2008, May). *The influence of variables on school report cards regarding the passing rates for students taking the high school proficiency assessment (HSPA) in New Jersey comprehensive high schools ProQuest LLC.* Retrieved from http://gradworks.umi.com/33/19/3319531.html

Kandel, D. B., & Lesser, G. S. (1969). Parental and peer influences on educational plans of adolescents. *American Sociological Review*, 34(2), 213-223.

Klebanov, P. K., Brooks-Gunn, J., & Duncan, G. J. (1994). Does neighborhood and family poverty affect mothers' parenting, mental health, and social support? *Journal of Marriage and Family, 56*, 441.

Kohn, M. L. (1963). Social class and parent-child relationships: An interpretation. *American Journal of Sociology, 68*, 471.

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher,*

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*(2), 4-16.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*(12), 1181.

Luster, T., Rhodes, K., & Hass, B. (1989). The relation between parental values and parenting behavior: A test of the Kohn hypothesis. *Journal of Marriage and Family, 51*, 139.

Madaus, G. F., Airasian, P. W., & Kellaghan, T. (1980). *School effectiveness: A reassessment of the evidence*. New York, NY: McGraw-Hill.

Marzano, R. J. (2000). *A new era of school reform: Going where the research takes us.* Aurora, Colorado: Mid-continent Research for Education and Learning.

Maylone, N. (2002). *The relationship of socioeconomic factors and district scores on the Michigan educational assessment program tests: An analysis.* (Unpublished doctoral dissertation). Eastern Michigan University,

Mehrens, W. A. (1998). Consequences of assessment: What is the evidence? *Education Policy Analysis Archives, 6*(13)

Meisels, S. J., & Shonkoff, J. P. (2000). *Early childhood intervention: A continuing evolution. in Shonkoff, J.P., & Meisels, S.J. (eds.) Handbook of early childhood intervnetions (2nd ed.)*. New York, NY: Cambridge University Press.

Messick, S. (1988). *The once and future issues of validity: Assessing the meaning and consequences of measurement*. Hillsdale, NJ: Lawrence Erlbaum Associaties.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp.13-103) New York: American Council on Education/Macmillan Publishing.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749.

Mintrop, H., & Sunderman, G. L. (2009). Predictable failure of federal sanctions-driven accountability for school improvement and why we may retain it anyway. American Educational Research Association. *38*(5), 353-364. Retrieved from http://edr.sagepub.com/content/38/5/353

Nagaoka, J., & Roderick, M. (2005). Retention under Chicago's high-stakes testing program: Helpful harmful, or harmless*? Educational Evaluation and Policy Analysis,* 27(4), 309-340.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, D.C.: U.S. Government Printing Office.

National Council on Measurement in Education. (1995). *Code of professional responsibilities in educational measurement*. Washington, D.C.: National Council on Measurement in Education.

National Defense Education Act (P.L. 85-864; 72 Stat. 1580), (1958).

National Education Association of the United States, Commission on the Reorganization of Secondary Education. (1918). *Cardinal principles of secondary edcuation: A report of the commission on the reorganization of secondary education, appointed by the National Education Association*, No. 31-36. Washington, D.C.: U.S. Government Printing Office.

National Governors Association. (2009). *Common core state standards initiative.* Washington, D.C.: NGA. Retrieved from www.corestandards.org

New Jersey Department of Education. (2009a). *Grades 3 and 4 New Jersey assessment of skills and knowledge: Technical report.* Trenton, NJ: New Jersey Department of Education.

New Jersey Department of Education. (2009b). *NJ ASK 2009 score interpretation manual.* Trenton, NJ: New Jersey Department of Education.

Olson, D. R. (2004). The triumph of hope over experience in the search for what works: A response to Slavin. *Educational Researcher,* January(33), 24.

Parker, F. W. (1894). *Talks on pedagogics*. New York, NY: E.L. Kellog.

Pauli, W. J. (1960). In praise of Bagley. *The Clearing House, 34*(6), 345.

Payne, K. J., & Biddle, B. J. (1999). Poor school funding, child poverty and mathematics achievement. *Educational Researcher, 28*(6), 4-13.

Pereira, M. A. (2011). *The influence of student and school variables on student performance on the New Jersey assessment of skills and knowledge in grade 8. Seton Hall University.* Retrieved from http://domapp01.shuedu/depts/uc/apps/library repository.nsf/resourceid/AF5C74EF331500F58525784E005100E4$File/Pereira_ Maria_A_Doctorate.pdf?Open

Peterson, J., & Zill, N. (1986). Marital disruption, parent-child relationships, and behavior problems. *Journal of Marriage and Family, 48*(2), 295-307.

Plug, E., & Vijverberg, W. (2005). Does family income matter for schooling outcomes? Using adoptees as a natural experiment. *The Economic Journal, 115*(506), 879-906.

Ramey, C. T., & Ramey, S. L. (1998). Early intervention and early experience. *American Psychologist, 53*(2), 109-120.

Reinard, J. C. (2006). *Communication research statistics*. Thousand Oaks, CA: Sage.

Roscigno, V. J., & Ainsworth-Darnell, J. J. (1999). Race, cultural capital, and educational resources: Persistent inequalities and achievement returns. *Sociology of Education, 72*(3), 158-178.

Rustique-Forrester, E. (2005). Accountability and the pressure to exclude: A cautionary tale from England. *Education Policy Analysis Archives, 13*(26) Retrieved from http://epaa.asu.edu/epaa/v13n26

Shonkoff, J. P., & Meisels, S. J. (2003). *Handbook of early childhood intervention*. New York, NY: Cambridge University Press.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417-453.

Sizer, T. R. (1984). *Horace's compromise: The dilemma of the American high school*. New York, NY: Houghton Mifflin Company.

Skinner, B. F. (1953). *Science and human behavior*. New York, NY: The Free Press.

Smith, J. R., Brooks-Gunn, J., & Klebanov, P. K. (1997). Consequences of living in poverty for young children's cognitive and verbal ability and early school achievement. In G. J. Duncan, & Brooks-Gunn, J. (Eds.) *Consequences of growing up poor*, pp. 55-68. New York, NY: Russell Sage Foundation.

Sobol, Thomas (2003). A smarter mind than mine takes on NCLB. Retrieved from http://practicaltheory.org/serendipity/index.php?/archives/757-A-Smarter-Mind-than-Mine-Takes-on-NCLB.html

Tanner, D. l., & Tanner, L. (2007). *Curriculum development: Theory into practice* (4th Ed.). Upper Saddle River, NJ: Pearson Merrill Prentice Hall.

Task Force on Education for Economic Growth. (1983). *Action for excellence: A comprehensive plan to improve our nation's schools*. Denver, CO: Education Commission of the States.

Taylor, F. W. (1911). *The principles of scientific management*. Norwood, MA: Harvard University Press.

Thorndike, E. L. (1911). *Animal intelligence: Experimental studies*. New Brunswick, NJ: Transaction Publishers.

Tienken, C. H. (2008(a)). A descriptive study of the technical characteristics of the results from the New Jersey assessments of skills and knowledge in grades 3,4, and 8. *New Jersey Journal of Supervision and Curriculum Development, 52*, 46-61.

Tienken, C. H. (2008b). A descriptive study of the technical characteristics of the results from the New Jersey assessment of skills and knowledge in grades 3,4, and 8. *New Jersey Journal of Supervision and Curriculum Development, 53*, 47.

Tienken, C. H. (2011). *Structured inequity: The intersection of socioeconomic status and the standard error of measurement of state mandated high school test results*. (pp. 257-271). NCPEA Yearbook. Lancaster, PA: Proactive Publications.

Tienken, C.H., & Rodriguez, O. (2010). The error of state mandated high school exams. *Academic Exchange Quarterly, 14*(2), 50-55.

Tomlinson, S. (1997). Edward Lee Thorndike and John Dewey on the science of education. *Oxford Review of Education,* 23(3), 365-383.

Weinberg, B. A. (2001). An incentive model of the effect of parental income on children. *Journal of Political Economy, 109*(2), 266.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*(3), 461-481.

White, S. B., Reynolds, P. D., Thomas, M. M., & Gitzlaff, J. (1993). Socioeconomic status and achievement revisited. *Urban Education, 28*(3), 328-343.

Zilversmit, A. (1993). *Changing schools: Progressive education theory and practice, 1930-1960*. Chicago, IL: University of Chicago.