Fall 2012

The Influence of Inclusion on the Academic Performance of General Education Students on the New Jersey Assessment of Skills and Knowledge in Grades 6, 7, and 8

Christie M. Robinson
Seton Hall University

The Influence of Inclusion on the Academic Performance of General Education Students on the New Jersey Assessment of Skills and Knowledge in Grades 6, 7, and 8

Christie M. Robinson

Dissertation Committee

Anthony Colella, Ph.D., Mentor
Christopher Tienken, Ed.D., Committee Member
Gerard Babo, Ed.D., Committee Member
Michele Wilson-Kamens, Ed.D., Committee Member

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

Seton Hall University

October 2012

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

APPROVAL FOR SUCCESSFUL DEFENSE

Doctoral Candidate, **Christie M. Robinson,** has successfully defended and made the required modifications to the text of the doctoral dissertation for the **Ed.D.** during this **Fall Semester 2012**.

DISSERTATION COMMITTEE
(please sign and date beside your name)

Mentor:
Dr. Anthony Colella 10/17/12

Committee Member:
Dr. Christopher Tieken 10/17/12

Committee Member:
Dr. Gerard Babo 10/17/12

External Reader:
Dr. Michele Kamens

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 impact of student variables (student socioeconomic status, race, attendance, and gender) and the school variable of placement in an inclusion setting on the academic achievement of general education students in grades 6, 7, and 8 in an urban school district. Academic achievement was defined as a general education student's performance in mathematics and language arts and literacy as measured by the New Jersey Assessment of Skills and Knowledge (NJ ASK) annual state test. Analyses were conducted using multiple regression models. The sample was comprised of approximately 1200 students enrolled in grades 6, 7. and 8 in a New Jersey urban district's middle schools during the years 2010-2011. Results of this study indicated that placement in an inclusion classroom did have a statistically significant impact on the NJ ASK scores of non-disabled students; therefore, further research needs to be conducted in the area of inclusion in order to determine why inclusion is having a negative impact on the academic achievement of non-disabled students.

Keywords: special education, inclusion, general education, academic achievement

Acknowledgments

To my dearest husband, Rob, my husband and biggest fan. I do not know what I would have done without you through this process. The nights when I would cry and want to give up, you kept my spirits high and had more confidence in my abilities than I had in myself. I appreciate all the sacrifices and dinners that you made for me while I completed this dissertation. I love you and thank you for being my partner through it all.

To my parents, Rosemary and Mike, and my brother, Mikey, this would not have been possible without all of your love and moral support. You always instilled in me that I could do anything that I put my mind to, and look where I am at 29 years old. I could not have asked for a more caring and amazing family. I am a lucky girl because you have given me the opportunity to pursue my dreams and be all that I can be. Mom, I am the luckiest girl in the world to have a mother who is such an amazing mom and principal. I have achieved greatness in my career because you have been an extraordinary mentor and role model. Dad, I am so lucky that I inherited your ability to relax and laugh because otherwise I would not have survived this dissertation process. To my little brother, Mikey, I beat you! I am so proud of you and all that you have accomplished; you are the best brother that I could ask for. Thank you for all your support and love.

Dr. Tienken, Dr. Colella, and Dr. Babo, I am so fortunate to have worked with such supportive, knowledgeable, and caring dissertation committee members. If it were not for your patience, ceaseless encouragement, positive notes, and emails, I would not have produced a

study that I am so proud of. Thank you for all of your help, support, and feedback. I am a better and more educated person because of all of you.

Dr. Kamens, I do not know how to thank you. You are the reason that I am at this point in my education and career. You have been one of the most inspirational teachers that I have ever had. Your constant encouragement through my master's degree and support throughout my doctoral studies has been unprecedented. Thank you for always being there for me and for introducing me to proper writing techniques through red marks on my assignments. Throughout our career as educators, we always hope that we have impacted a life. I do not hesitate to say that because of you I am Christie M. Robinson, Ed. D. I hope that I will be your favorite success story.

Dedication

This work is dedicated in loving memory of Dorothy Robinson, Gerald Lombardo, and Carmela Lubak, my dearest relatives, who left this world way too soon and who would have to loved to be here with me to celebrate and share in this moment. You are always in my thoughts.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter I

INTRODUCTION

The new decade of 2010 began with a shift in paradigm with regard to special education policy. Beginning in 1975, with the adoption of the Education for All Handicapped Children Act and continuing with its revision to the present day legislation entitled Individuals with Disabilities Education Improvement Act (IDEIA), policy in special education has shifted from a focus of isolation and small group instruction to a movement of inclusion and access to an education in the "least restrictive environment" for children classified as special education (Baker & Zigmond, 1995; Faas, 1980). Policymakers have prompted the restructuring of classrooms by "the abandonment of pull-out and the return of students to the general education setting while delivering whatever instruction is needed within the confines of the general education classroom" (Baker et al., 1995, p. 172). In other words, a plethora of time and money has been invested in special education reform as strides are being made in adding "new teachers and supports to inclusive settings, allowing for the successful inclusion of students with disabilities" (Lewis, 2002, p. 114).

Along with the revision of IDEIA, President George W. Bush implemented a federal education policy in 2001 entitled *No Child Left Behind (NCLB)*. The accountability and testing mandates required by writers of NCLB reinforced that the focus in education remain on research and evidence-based practices, standardized testing, and federal control. January 2010 marked the eighth anniversary of the *No Child Left Behind legislation,* and the journey from its inception in 2001 has been "fraught with controversy as the federal government has assumed a broader and more forceful role in elementary and secondary education" (Dietz, p. 16). Within this control,

attempts have been made to improve the education system nationwide including the following ideas: standards-based reform, national standards, annual testing in reading and math, public reporting of test results by school, disaggregated by specific student groups, and approaches for determining improvement and identifying low-performing schools (Dietz, p. 26).

As federal education policy continues to drive state education systems, we find ourselves at a critical juncture. Federal policy now reaches into every school district and classroom in the country. States routinely submit plans to the federal government about who will teach in their schools and when children will be tested. Teachers are fired and schools closed based on federal policy (Dietz, p. 32). In other words, school district administrators are under constant scrutiny to achieve mandated test scores defining their school as "proficient" rather than being labeled "failing."

Along with this struggle to maintain vigorous accountability standards, another mandate of the No Child Left Behind legislation legally requires "expanding the range of accommodations available to students with disabilities to improve the rates of inclusion of such students in the general education population" (Public Law 107-110: 115 STAT. 1874). Although this idea was introduced with the development of the Individuals with Disabilities Act (IDEA), it still causes much controversy among school officials. No longer are students with disabilities exempt from meeting "typical benchmarks." According to the mandates of NCLB, students with disabilities must be tested on grade level with their non-disabled peers and provided the use of necessary modifications. Because students classified as special education are deemed a mandatory "sub group," a school's success and/or failure depends on the outcome of the standardized assessment scores of the special education population.

Before the landmark case of *Brown v. Board of Education (1954),* in which the Supreme Court of the United States mandated to racially desegregate schools, individuals with disabilities were forced to be segregated from their age appropriate non-disabled peers, locked away in state institutions that were of inhumane and unsanitary conditions. Consequently, children with "mental retardation" were "separated" and given no opportunities to learn or be exposed to daily life skills, leaving them dependent on others for survival.

Advocacy for a change in law and policy with regard to special education began in the early 1970's with an eyewitness investigation conducted by Geraldo Rivera, a newscaster and journalist for the State of New York. Rivera's investigation would expose the world to the unsanitary and repulsive conditions at Willowbrook State School and record a shockingly truthful documentary. As a result, a class-action lawsuit was filed against the State of New York in Federal Court on March 17, 1972. A settlement in the case was reached on May 5, 1975, mandating reforms at the site. Unfortunately, several years would elapse before all of the violations were corrected. As a result, the publicity generated by the case was a major contributing factor to the passage of a federal law entitled the *Civil Rights of Institutionalized Persons Act of 1980*, granting protection to those in state or local facilities against harmful and unsanitary living conditions. (LaMorte, 2008, p. 108).

Challenges to existing law in special education began even before the incident at Willowbrook with the Elementary and Secondary Education Act of 1965. This act continued to be modified pending outcomes of new court rulings, but this legislation did not provide for inclusion or a "free and appropriate education" (FAPE). However, in 1972 two cases, *Pennsylvania Association for Retarded Children v. Pennsylvania (1972) and Mills v. Board of*

Education of the District of Columbia (1972), resulted in making significant progress in advocating for educational rights for children with disabilities. The decisions of these two court cases provided "disabled students ages 6 to 21 access to a free public education in an appropriate placement (regular education or special education classes) depending on individual needs" (LaMorte, 2008, p. 85). Additionally, Mills clarified the initial ruling to include the word *adequate* meaning that the education received by each child met his or her individual developmental learning needs (2008).

The history of IDEA began in 1975 with the passage of the Education for All Handicapped Children Act of 1975, which required public schools accepting federal funds to provide equal access to the educational curriculum for children with disabilities as well as requiring the school district to evaluate each student with a disability and, with the help of the parents, create an educational plan that would be similar to that of his/her non-disabled peers. Overall, the Education for All Handicapped Children Act was created to set the foundation for special education law in four distinct areas: (1) to ensure that special education services were available in public schools for any child requiring them, (2) to require school districts to make fair and appropriate decisions about special education services, (3) to create a system of determining requirements for special education services, and (4) to provide federal funding to public schools strictly for the education of students with special needs.

Throughout the next two decades, the Education for All Handicapped Children Act would undergo many revisions with the addition of education for infants and toddlers (P.L. 99-457), and attorney's fees (P.L. 99-373). One of the biggest changes to Public Law 94-142 would occur during 1990, when it would be renamed the Individuals with Disabilities Act (IDEA) and provide

for the following services by law: transition services, terminology changes (handicapped children to children with disabilities), and additional information regarding funding, eligibility, and Individualized Education Plans (IEPs) (LaMorte, 2008, p. 67).

IDEA was amended by Congress in 1997 and reauthorized in 2004 as the Individuals with Disabilities Education Improvement Act. The language in IDEIA was further modified to include specific mandates to public schools which included the following: public schools must provide a free and appropriate education "to the maximum extent appropriate, children with disabilities, including children in public or private institutions and other care facilities, are educated with children who are not disabled, and children with disabilities must be educated in their least restrictive environment to the maximum extent possible" (IDEIA, p. 6).

These laws translate from federal legislation to mandate by the state. In New Jersey, officials have transcribed the federal laws into the New Jersey School Administrators Code, which can also be referred to as N.J.S.A., Title 6A, Chapter 14, and which specifically refers to the area of special education. New Jersey Code 6A:14-1.2 outlines that each district board of education shall have policies, procedures, and programs approved by the Department of Education through the county office of education that are in effect to ensure the following mandates. The following code mandates outlined by the New Jersey Code are clearly drafted and connected to the federal mandate of IDEA (94-142) as well as the No Child Left Behind (NCLB) legislation:

1. A free appropriate public education according to N.J.A.C. 6A:14-1.1 is available to all students with disabilities between the ages of 3-21, including students with disabilities that have been suspended or expelled from school.

2. Each board of education is responsible for providing a system of free, appropriate special education and related services to students with disabilities, ages 3-21, which shall be provided at public expense, under public supervision, with no charge to the parent (N.J.A.C. 6A:14-1.2).

As state standardized tests continue to be the primary indicator to measure student achievement, emphasis has been placed on the benefits of education in an inclusive environment for students with disabilities. Many studies have shown both the academic and social benefits of the inclusive environment on student with disabilities. Baker, Wang, and Walberg (1994) cite evidence that "students with disabilities educated in regular classes do better academically and socially than comparable students with disabilities in noninclusive settings" (p. 34). Additionally, qualitative findings by Idol (2006) strongly support the practice of "including students with special education challenges in general education programs" due to social and academic gains (p. 82).

Although researchers have provided data showing the benefits of an inclusive setting for students with disabilities, the influence of the inclusive classroom on the achievement of students without disabilities yields a variety of mixed results and conclusions. In fact, it would be useful for school administrators to know how the New Jersey Assessment of Skills and Knowledge (NJASK) scores of students without disabilities are influenced when they are placed in an inclusive environment.

A study conducted by Afroditi Kalambouka (2007) presents evidence supporting that "there are no adverse effects on pupils without disabilities in mainstream schools, results warranted either a neutral (no difference) or a positive gain in academic achievement" (p. 362).

There is also research that shows that general education students' reading scores were not significantly affected by being placed in an inclusive setting (Gandhi, 2007; Huber, Rosenfeld, & Fiorello, 2001). However, additional research has shown that "direct teaching and coaching became the responsibility of an 'assigned study buddy' (a general education student assigned to a classified student, an added responsibility not a part of a typical general education structure" (Baker & Zigmond, 1995, p. 176).

Additional evidence exists indicating that positive or negative academic results are based on the "type" of classified student placed in an inclusive setting (Kalambouka, 2007, p. 376). Kalambouka found that placing students with behavioral and emotional needs in inclusive settings yielded negative outcomes compared to students with other documented disabilities (2007). Due to a lack of empirical evidence, most researchers express caution in drawing conclusions based on their findings. According to Baker and Kalambouka, the limited number of studies and the vagueness in the definition of "inclusion" all become relevant when planning and making decisions on policy and practice.

Statement of the Problem

Federal policies such as NCLB and IDEA require all students in special education to be educated in their least restrictive environment to the maximum extent possible, which in some cases includes placement in an "inclusive setting." In order to be deemed both "proficient (meeting the Annual Yearly Progress requirements of NCLB) and simultaneously remaining in compliance with legal mandates of providing a free and appropriate education (FAPE), principals and school leaders throughout the field of education struggle with creating programs that foster

the "inclusive environment" while improving the academic achievement of all students to meet the established AYP mandates.

The outcome of high-stakes, standards-driven state assessments determines whether schools are categorized as making adequately yearly progress (AYP) toward the 2014 target of 100 percent proficiency for all students or deemed in need of improvement (NCLB, p. 110). Consequently, student performance on high-stakes standardized assessments has become a primary indicator of success or lack of success for students, teachers, administrators, schools, and school systems. As cases are made for and against inclusion, principals debate where to place general education students--in a traditional classroom structure or an inclusive setting. Many continue to question whether the inclusive environment will foster and/or improve general education student performance on state standardized tests.

The majority of the quantitative evidence suggests there are academic and emotional benefits for students with disabilities; however, a smaller body of research addresses the effect/ benefits for students without disabilities. The body of empirical research that does exist concerning the influence of inclusion on non-disabled student's academic achievement has resulted in mixed outcomes (Daniel & King, 1997; Huber, Rosenfeld, & Fiorello, 2001; Hunt, Staub, Alwell, & Goetz, 1994; Kalambouka et al., 2007; Manset & Semmel, 1997; Saint-Laurent et al., 2002; Sharpe, York, & Knight, 1994; Staub & Peck, 1995). Limits to research in this area also tend to group general education students based on a "one size fits all" model, disregarding variables that may also influence or hinder their academic achievement (Daniel & King, 1997). Research identifies the variables that influence student achievement as follows: student attendance, socio-economic status, eligibility for free lunch, and race/ethnicity.

The problem rests with the lack of empirical quantitative evidence explaining the influence of inclusion on the student achievement of general education students. In other words, policymakers have continued to focus on the benefits of inclusion for students with disabilities and neglected to consider that the general education population of students is not homogeneous in ability; therefore, these policies and placements in an inclusion setting could influence their academic achievement. This study will yield additional insight into the effects of placement in an inclusion setting on the academic achievement of specific subgroups within the general education population, thus fostering new knowledge in determining the best placement for students within the general student population.

Purpose of the Study

The purpose of this study is to determine whether placement in an inclusive setting affects the academic achievement of general education students on the Language Arts Literacy and Mathematics section of the New Jersey Assessment of Skills and Knowledge (NJ ASK), Grades 6, 7, and 8. Additionally, this study aims to examine specific models including the independent variables of inclusive setting, non-inclusive setting, student attendance, and eligibility for free and reduced lunch that, paired with placement in an inclusive/non-inclusive setting, may result in an effect on the dependent variable of student achievement on the NJ ASK Grades 6-8. As emphasis continues to be placed on the outcomes of the NJ ASK scores to determine school accountability, district administrators must consider the needs of all students. This study aims to produce research-based evidence to continue the development of policy and create the opportunity for educated, data-driven decisions that benefit all students. As a result,

school professionals will choose an instructional program that meets individual students needs and that will maximize both learning and student achievement.

Research Questions

The following research questions guided this study:

Research Question 1: What is the influence of placement in the inclusive setting on the performance of non-disabled students in the area of mathematics as measured by the NJ ASK when controlling for student mutable variables at Grades 6, 7, and 8?

Research Question 2: What is the influence of placement in the inclusive setting on the performance for non-disabled students in Grades 6, 7, and 8 in the area of language arts and literacy when controlling for student mutable variables as measured by the NJ ASK?

Research Question 3: How well does placement in an inclusion classroom, student gender, attendance in class, race, and SES predict and/or influence student overall academic performance as measured by the NJ ASK 6, 7, and 8?

Null Hypothesis:

H_o^1: Placement in an inclusive setting has no influence on the performance of non-disabled students in Grades 6-8 in the area of mathematics as measured by the NJ ASK.

H_o^2: Placement in an inclusive setting has no influence on the performance of non-disabled students in Grades 6-8 in the area of language arts and literacy as measured by the NJ ASK.

H_o^3: Placement in an inclusion classroom, student gender, attendance in class, race, and SES have no influence on the performance of non-disabled students in the areas of language arts and mathematics as measured by the NJ ASK.

Significance of the Study

In recent years, inclusion of students with disabilities into regular education classrooms and creating inclusive schools has become a priority of the education movement in order to provide a free and appropriate education (FAPE) as well as ensure equal learning opportunities for students classified as special needs. The problem lies in the fact that "many schools continue to resist the pressure to become more inclusive because they are concerned that to do so will have a negative influence on the academic progress of other pupils and/or will lower academic standards" (Florian, Rouse, Black-Hawkins, & Jull, 2004, p. 115).

With that said, results of this study are significant in that they will benefit school administrators, educators, parents, and researchers in filling in the research gap that exists in determining whether placing general education students in an inclusive setting has an influence on their student achievement as measured by the state standardized tests. Klingner, Vaughn, Hughes, Schumm, & Elbaum (1998) explain plainly that "although discussions of the pros and cons of inclusion are likely to continue, many recognize that what is missing is empirical evidence that documents the effects of inclusion, particularly for students without learning disabilities (p.153). Additionally, it will provide insight into the variability of the general education population of students, which is commonly looked at as homogeneous, and render valuable data to aid school leaders and parents in deciding whether an inclusion placement is in the best interest of each general education child. With more information, we can examine inclusion further and determine what is indeed best placement and best academic practice for each individual child so that he or she may reach academic potential. It is imperative for school

stakeholders and policymakers to look at inclusion from multiple perspectives, not only the view of the special needs student.

Limitations

The literature that was analyzed serves to add to the current body of research on the inclusive classroom and its influence on the student achievement of non-disabled students. Caution must be exercised when making generalizations based on the findings of this study because limitations apply to this quantitative research design.

First, Harwell and Lebeau (2010) found that of all the SES measures available, free and reduced lunch eligibility was the most likely to be used because it provides easy access, is inexpensive, and requires minimal responses from participants in the sample; however, it is not valid as an indicator of access to household resources. With that said, the authors recommend that an important practice for education researchers is to "adopt and carefully describe what SES is intended to represent in his or her study and to provide a clear rationale for selecting that measure of SES" (p. 126). It is also important to disclose the limitations of choosing to use free and reduced lunch (FRL) as an input variable. The limitations of using FRL as an SES variable include the following:

1. Eligibility for FRL is a poor measure of a student's access to economic resources because FRL is not strictly based on federal poverty guidelines. There are other economic factors that must be taken into account in order to obtain a true picture of a family's SES.

2. Students can directly qualify for FRL in ways that do not directly coincide with household income such as living in a foster home.

3. Students can be incorrectly certified as eligible or not eligible based on a variety of aspects. A recent study has shown that 17% of students certified as eligible for FRL should not have been, and 8% certified as ineligible for reasons such as late applications and/or administrative error were in fact eligible. This study does use the variable of eligibility for free lunch as a measure of SES.

There are other limitations that should be noted as well. The students in this study were examined as a population of students who were placed in an inclusive setting. Individual class makeup was not known. Assumptions were made based on the mean test scores of groups of general education students both in and not in inclusive settings. Because the academic makeup of each inclusion class was unknown, it is hard to decide whether it is the variable of inclusion having the influence or the fact that all of these students grouped together with the students in special education is creating the negative impact on academic achievement. This limitation, although disclosed, was attempted to be rectified by using prior achievement as an independent variable as well as conducting hierarchical linear regression to determine confounding and/or suppressor variables.

It is also important to note that the definition of an "inclusive classroom" can be considered a limitation because we do not know the quality or context of practice within these classrooms labeled as "inclusive." For example, do the teachers assigned to the inclusive classrooms use models and structures of current practice, or is the special education teacher there only to support students with disabilities? These are issues related to inclusive practice that would have impact on the results of this study.

Also, this study takes place in two urban middle schools, Grades 6-8, in a low socio-economic school district. Results may not be generalizable to suburban or higher socio-economically based school districts. Also, this study does not include specific credentials of teachers who have worked with the students in current and past years as well as specifics regarding classroom makeup.

The correlational design does not allow for causal comparisons. The design can help to identify relationships among variables.

Delimitations

Delimitations for the study were as follows:

1. New Jersey Ask Scores were collected for both the 2009-2010 and 2010-2011 school years in both language arts and mathematics.

2. The study focused on one district in the B district factor group.

3. The research incorporated Grades 6, 7, and 8.

4. Student prior achievement was controlled for.

Definition of Terms

Academic Achievement -- measured by individual student scores on the New Jersey Assessment of Skills and Knowledge (NJ ASK) for Grades 6-8 in language arts and literacy. The NJ ASK tests students' knowledge and achievement in the New Jersey Core Curriculum Content Standards.

Adequate Yearly Progress (AYP) is the target set by each state, based on meeting the No Child Left Behind Act's overall goal that all students be proficient in reading and math curriculum standards by 2014.

New Jersey Assessment of Skills and Knowledge (NJ ASK) -- a criterion-referenced standards-based standardized test designed specifically to measure the degree to which all students in Grades 3-11 have attained proficiency in the New Jersey Core Curriculum Content Standards (CCCS) in Language Arts Literacy (LAL), mathematics, and science (excluded in Grade 3).

Classroom Setting – refers to student placement; students are either placed in an inclusive setting which includes students classified as in need of special education or a general education setting made up entirely of students not deemed in need of special education services.

Co-teaching – also known as team teaching, cooperative teaching, or collaborative teaching, is the process by which a general educator and a special educator teach together in an inclusive classroom (Stuart et al., 2006).

Mainstreaming – a term that typically refers to the placement of a child with special developmental, physical, emotional, or educational deficiencies or challenges into a regular classroom setting for part or all of the school day, with the long-term goal of helping the child make a gradual adjustment into as many aspects of normal life as possible, so that the child can become a functioning member of society to whatever extent he or she is able (Hardman, Drew, & Egan, 2002, p. 38).

Inclusion – educating disabled students with their non-disabled age appropriate peers to the maximum extent appropriate with appropriate aids, modifications, and supports in the general education classroom in the school the student would attend if he or she did not have a disability

General Education -- integrated learning experiences structured across subject disciplines to provide the set of skills and knowledge needed to function in society (Sternberg & Williams, 2002, p. 152)

Special Education -- instruction that is specially designed to meet the unique needs of a child with a disability typically associated with an Individual Education Plan. Specifically, education that is developed to address an individual child's needs that stem from his or her disability.

Student with a Disability – refers to a student who has been classified with a disability under IDEIA (P.L. 108-446) and New Jersey Administrative Code 6A:14-3.5 into categories as follows: auditory impaired, autistic, cognitively impaired, communication impaired, emotionally disturbed, multiply disabled, deaf/blindness, orthopedically impaired, other health impaired, social maladjustment, specific learning disability, traumatic brain injury, visually impaired, or Section 504 of the Rehabilitation Act of 1973 (Section 504, 29 U.S.C. § 794); specifically, a physical or mental disability that substantially influences a major life activity (e.g., walking, hearing, breathing, learning).

Student without a Disability – also known as a general education student, refers to a student who has not been classified with a disability under IDEA, New Jersey Administrative Code 6A:14-3.5 or Section 504 of the Rehabilitation Act of 1973.

Inclusive Class – general education setting where disabled and non-disabled students are educated together.

Non-inclusive or General Education Class -- educational setting that is comprised of non-disabled students

Special Education Class -- the education setting where disabled students are educated; i.e., resource center, self-contained classes (multiply disabled, autism, behavioral disabilities, learning and language disability).

Least Restrictive Environment – refers to the provision in the IDEA mandates requiring to the maximum extent appropriate that students with disabilities ages 3 through 21 are educated with non-disabled children and participate in nonacademic and extracurricular activities with non-disabled children. (IDEA, 20 U.S.C. § 1412 [a][5])

Free and Reduced Lunch -- Students in New Jersey are eligible for free and reduced lunch if the household income of their parents meet the following criteria:

Table 1

Eligibility for Free and Reduced Lunch

Your children may qualify for free or reduced price meals if your household income falls within the limits on this chart.

FEDERAL INCOME CHART For School Year 2007-2008			
Household size	Yearly	Monthly	Weekly
1	18,889	1,575	364
2	25,327	2,111	488
3	31,765	2,648	611
4	38,203	3,184	735
5	44,641	3,721	859
6	51,079	4,257	983
7	57,517	4,794	1,107
8	63,955	5,330	1,230
For each additional person, add:	6,438	537	124

Pending the approval of the application, a student will receive lunch at a reduced rate at no cost to them.

Student Attendance -- As per the district attendance policy, "The Board of Education requires the pupils enrolled in the schools of this district attend school regularly in accordance with

the laws of the State. The educational program offered by this district is predicated on the presence of the pupil and requires continuity of instruction and classroom participation. The regular contact of pupils with one another in the classroom and their participation in a well-planned instructional activity under the tutelage of a competent teacher are vital to this purpose" (Linden Board of Education Policy on Attendance, 2011, p. 2). For the purposes of this study, students will be categorized by days absent.

Basic Skills Instruction (BSI) -- Students are considered eligible for basic skills services when they are identified as "not proficient" on the NJ ASK in both math and language arts. They students are identified by teachers as "at-risk"; however, they are not eligible for special education services.

Chapter II

REVIEW OF THE LITERATURE

Introduction

Although current policy continues to drive the inclusive movement, by providing students in special education access to the general education population and curriculum, little consideration has been given to whether an inclusive setting is an appropriate placement for general education students and whether or not this placement has a positive or negative influence on their academic achievement.

As the accountability mandates increase annually, requiring a greater percentile of students to be deemed "proficient," research is needed to help school leaders make informed decisions on the best placement for all students, not only those with special needs. Because of the limited pool of research providing valid evidence on the influence of the inclusive setting on general education student achievement, it is often overlooked that the general education population is not "one size fits all." In fact, many studies continue to provide evidence that there are variables within the general education population that have a negative influence on student achievement, such as SES documented by schools as eligibility for free and reduced lunch, student attendance, race, and gender. It is possible that the combination of these variables with placement in an inclusion setting could play a role in positively or negatively influencing the academic achievement of general education students. The results of this study are imperative to enhance the small body of literature as the number of students in special education serviced within the public school system grows nationwide. While specific provisions in the law of

IDEA develop more access to the general curriculum for students in special education, research on inclusive practices is necessary to understand its effects and the barriers to improving overall student academic achievement and school accountability for all students (28th Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act, 2006.

The purpose of this study is to determine the extent to which placement in an inclusive setting influences the academic achievement of non-disabled students. Specifically, a statistical analysis will be conducted analyzing the independent/predictor variables of placement in an inclusive setting--student eligibility for free and reduced lunch, student attendance, race, and gender--and their influence on the dependent variable of academic achievement as measured by the Grade 6-8 New Jersey Assessment of Skills and Knowledge for Language Arts and Literacy (LAL) and Mathematics.

The review of the literature is divided into nine sections, including the following:

1. Historical Development of Inclusion,
2. Inclusion and Empirical Studies (
3. Empirical Studies on Academic Outcomes for Non-disabled Elementary Students in Inclusive Settings
4. Student Eligibility for Free and Reduced lunch and Its Influence on Student Achievement
5. Student Mobility and Its Influence on Student Achievement
6. A Review of the Excluded Variable of Student Gender
7. Summary

Literature Search Procedures

A thorough search was conducted to find all relevant literature that pertains to the topics addressed in this study. This included reviews of dissertations, relevant historical texts, and peer-reviewed research articles that meet the criteria outlined for the purposes of this study. Electronic publications were obtained through educational data bases including ERIC, EBSCO Host, Academic Search Premier, Lexus Nexus, and Seton Hall Dissertation Abstracts. Search parameters (advanced search options) were used to ensure that the literature appeared in peer-reviewed journals. Research was not limited to educational data bases. General web-based searches were conducted employing the use of Google and Yahoo to access various professional websites such as the New Jersey Department of Education, ed.gov, and the Center for Education Policy, looking for additional information, articles and legislative information. Keywords used include the following: history and special education, inclusion, special education inclusion, influence of inclusion on non-disabled students, influence of inclusion on disabled students, academic achievement, academic performance, gender, gender gap, NJ ASK, socio-economic status, poverty and academic achievement, English as a Second Language (ESL), English language learner (ELL), student mobility, special education law, basic skills instruction (BSI), academic achievement in language arts and mathematics. The noted terms and phrases were inputted in a variety of ways--in combinations or individually to produce said research results used in this study.

Criteria for Research

Criteria for studies used in this literature review include the following:

1. The studies involved elementary (pre-K-5), and middle school (6-8) students in school in the United States as well as districts worldwide.

2. Studies available via worldwide locations had to be either translated into English or readily available in English.

3. Non-peer-reviewed literature was only referenced for historical and legal purposes.

4. Studies included used a quantitative methodology, thus reporting on empirically-based findings. Quantitative studies that made comparisons without control groups were excluded. Additionally, studies that employed only quantitative analysis/findings on pre- and post-test comparisons were excluded.

5. Because of the vast number of definitions of inclusion, all studies that at minimum defined inclusion as "students in special education learning in a classroom with regular education students" were included. With that said, all categories of students in special education were included which are but are not limited to the following: students with social and emotional disabilities, autism spectrum disorder, students with learning disabilities (SLD), behavioral disorders, communication delays, multiple disabilities, and physical and cognitive disabilities.

6. Studies addressing "student achievement" had to employ the use of a standardized assessment in either language arts or mathematics given to both the control and the test groups. Studies that used researcher-created assessments or software-based programs were excluded, as they are not an accurate measure of curriculum standards.

Historical Development of Inclusion

Many would assume that the inclusion of students with disabilities and the legal obligation for schools to educate all students with disabilities occurred throughout the history of education; however, this assumption, although commonly accepted, would be incorrect. In fact,

prior to the 1970s millions of children with disabilities were either refused enrollment or inadequately served by public schools (Martin, et. al., 1996). Only over the last twenty years has the government begun to provide equal rights for disabled persons and abandon placements in state run institutions, allowing individuals with disabilities access to the educational system and the workplace. Until the late 1970s, services provided to children with disabilities and their families were minimal and at the discretion of the local school districts who had the right to refuse to enroll any student they considered "uneducable." (Martin, et. al., 1996).

The first signs of providing minimal services to individuals with disabilities began in the mid 1800s. The government pledged money to create "asylums for the deaf, the dumb, and the blind" (Public Law 45-186) (3/3/1879). From that point in historic America until the early 1950s the federal government was less than involved in the public school system until the landmark case of *Brown v. Board of Education* in 1954, which extended equal protection under the law to minorities. Once the decision from Brown mandated the racial desegregation of schools, advocates for individuals with disabilities began campaigning for desegregation for individuals with disabilities. Advocates demanded coordination of the federal government and educational institutions for children with disabilities, increased funding for the sole purpose of educating students with disabilities and "enforceable entitlement, eventually obtained through the courts" (1996). Dozens of cases were filed in courtrooms across the country; however, none of the decisions would influence individuals with disabilities until the early 1970s.

In an effort to quiet the legal disputes, education for individuals with disabilities was addressed again in 1958 when President Eisenhower signed a minor act (Public Law 85-926) as an amendment to the National Defense Education Act of 1958, providing financial support to

colleges and universities for training personnel in teaching children with mental retardation. The amendment was later expanded in 1963 to include the training of college teachers and researchers (Public Law 88-164). As members of the federal government continued to increase their role in education, they passed the Elementary and Secondary Act of 1965 (ESEA, P.L. 89-10), enacted to address the problem of inequity in education. Early legislation contributed to set the foundation for similar gains for individuals with disabilities.

Unfortunately, this did not satisfy advocates as they continued to lobby for an administrative unit within the U.S. Department of Education. In 1966, Congress mandated a Bureau for the Education of the Handicapped (BEH) under Title VI of ESEA (Elementary and Secondary Education Act). Known by the public as Title VI, this program would formulate the first Education of the Handicapped Act in 1970. It is important to note that this early legislation did not provide for "mainstreaming" or "free and appropriate education"; however, it did establish groundwork for future legislation concerning these issues (LaMorte, 2008).

During the pivotal years of the 1970s there was no state that served all of its children with disabilities. Many children were turned away while others were inappropriately placed in institutions or restrictive programs designed for more profoundly disabled individuals. Frustrated by these discriminatory practices, two critical court cases *Pennsylvania Association for Retarded Children (PARC) v. Commonwealth of Pennsylvania (1972) and Mills v. Board of Education of the District of Columbia (1972),* resulted in landmark decisions recognizing educational rights for children with disabilities.

In *Pennsylvania Association for Retarded Children (PARC) v. Commonwealth of Pennsylvania,* advocates contested a state law that "specifically allowed public schools to deny

services to children who had not attained a mental age of 5 years at the time that they would ordinarily enroll in first grade" (Martin, et al., 1996, p. 25). At the conclusion of the case, the decision rendered the agreement by state officials to provide "full access to a free public education to children with mental retardation up to age 21, a standard of appropriateness (an education appropriate to his or her learning capacities) and establishment of a clear preference for the least restrictive placement for each child" (*Pennsylvania Association for Retarded Children (PARC) v. Commonwealth of Pennsylvania, 1972, p. 2).*

Following the PARC decision, another suit was filed via *Mills v. Board of Education (1972).* During the time of Mills, it was determined that seven children between the ages of 8 and 16 were refused enrollment in the District of Columbia Public Schools due to the nature of their disabilities. Additionally, students were unjustly expelled due solely to their disability. After a thorough investigation, school district officials admitted to not servicing 12,340 students with disabilities because of "budget constraints" (LaMorte, 2008). At the conclusion of the case, the D.C. court judges ruled that school districts were "constitutionally prohibited from denying students with disabilities an education due to inadequate resources because the "equal protection clause of the 14th Amendment would not allow the burden of insufficient funding to fall more heavily on children with disabilities than on other children" (*Mills v. Board of Education*, 1972, p.18).

The decisions from PARC and Mills were a turning point for both state and federal special education legislation. Most importantly, further investigation revealed that as in Mills, "3.5 million children with disabilities were not being provided an education that meets their needs, while over one million were not receiving an education at all" (U.S. Department of

Education, 1983, p. 12). In response to this injustice and in an effort to defray additional court cases, Congress created two public laws, The Rehabilitation Act at Section 504 (Public Law 93-112) in 1973, and the Education for All Handicapped Children Act (Public Law 94-142) in 1975.

The Rehabilitation Act at Section 504 "provided that any recipient of federal financial assistance must end discrimination in the offering of its services to persons with disabilities" (P.L. 93-112, 87 Stat. 355, p.106). This law applied to both state and local educational agencies. Ironically, the government provided no money/funding and no monitoring for Section 504 resulting in its being ineffective for about 20 years. Expansion of Public Law 93-112 did not arrive until 1990, when Congress passed the American's with Disabilities Act (ADA), which provided for additional rights for people with disabilities. ADA expanded non-discrimination practices into places of employment, public accommodations, transportation, and various methods of communication.

Two years after the creation of Public Law 93-112, saw the creation of Public Law 94-142, the Education for All Handicapped Children Act. This act, passed by Congress in 1975 and signed by President Gerald Ford, would require that all students with disabilities receive a free, appropriate public education and that school districts be provided funding to help with excess costs when developing special education programs. P.L. 94-142 required public schools accepting federal funds to provide equal access to the educational curriculum for children with disabilities as well as requiring the school district to evaluate each student with a disability, and with the help of the parents, create an educational plan that would be similar to that of his or her

non-disabled peers. Overall, P.L. 94-142 was created to set the foundation for special education law in four distinct areas:

1. To ensure that special education services were available in public schools for any child requiring them

2. To require school districts to make fair and appropriate decisions about special education services

3. To create a system of determining requirements for special education services

4. To provide federal funding to public schools strictly for the education of students with special needs.

President Ford expressed concerns that "the bill would be too expensive, would interfere with state responsibility, and would upset the balance of relationships between parents and local schools" (Martin et al., 1996, p. 36). President Ford's concerns were ignored as the act continued to be modified throughout the 1980s through the newly formed Department of Education's Office of Special Education and Rehabilitation Services and then was renamed in 1990 as P.L. 101-476, *The Individuals with Disabilities Education Act (IDEA*).

IDEA is the landmark legislation in special education. It provided the foundation of detailed legal mandates that must be followed in order for school districts to remain in compliance with the law. IDEA, "authorizes federal funding for special education and related services and for states that accept these funds, sets out principles under which special education and related services are to be provided" (Apling & Jones, 2002, p. 1). IDEA (20 U.S.C. 1400 et. Seq,) provides strict guidelines to ensure that each child with a disability receives a free,

appropriate public education (FAPE) (Apling & Jones, 2002, p. 2). The major principles of IDEA are as follows:

1. States and school districts make available a free and appropriate public education (FAPE) to all children with disabilities between the ages of 3 and 21. In order to provide these services, state departments of education and school district officials must identify, locate, and evaluate all children with disabilities, regardless of how severe, and determine which of these children are eligible for special education and related services.

2. IDEA requires that each child receiving services has an individualized education program (IEP) spelling out the specific special education and related services to be provided. These services must meet the child's individual needs, and the parent must be a partner in planning and overseeing the child's special education and related services as a member of the IEP team.

3. To the maximum extent appropriate, children with disabilities must be educated with children who are not disabled, and states and school districts must provide parents and guardians with procedures in order to appeal decisions--the right to a due process hearing, the right to appeal in federal district court, and the right to receive attorneys' fees (Apling & Jones, 2002, p. 4).

This legislation continued to shape and form special education law and policy over the next decade until its reauthorization by President Bush on December 3, 2004. IDEA P.L. 101-476 was reauthorized in 2004 as the Individuals with Disabilities Education Improvement Act (IDEIA, P.L. 108-446). Although IDEIA included the same principles as IDEA, it added

major revisions to the law, which coincided with NCLB of 2001. These revised mandates include but are not limited to the following:

1. An extensive definition of "highly qualified" special education teachers and a requirement that all special education teachers must be highly qualified

2. Extensive provisions aimed at ensuring that special education and related services provided for children with disabilities who are homeless or members of highly mobile populations (Child Find).

3. Authorization provided for states to use IDEA funding to establish and maintain "risk pools" to aid local educational agencies.

4. Modifications to requirements for parents who unilaterally place their children with disabilities in private schools

5. Revised state performance goals and requirements for children's participation in state and local assessments to align these requirements with those in the ESEA

6. Authority for education institutions to use their local IDEA grant for "early intervention services" aimed at reducing or eliminating the future need for special education for children with educational needs who do not qualify for IDEA.

7. Significant changes to parents' rights and procedural safeguards, including the addition of a "mediation hearing" prior to due process to try to resolve conflict/dispute and revised tests regarding manifestation determination. The school has the ability to place a child with a disability in an interim alternative education setting when a child has inflicted serious bodily injury on another person.

8. Changes in compliance monitoring with a focus on student performance.

9. Authority to extend service for infant and toddler services beyond the age of 2 (Apling & Jones, 2005, p. 2).

Research questions in this study address areas required by the IDEIA legislation, most importantly "Least Restrictive Environment (LRE)." To comply with LRE requirements, IDEIA mandates the following:

> ...to the maximum extent appropriate, children with disabilities, including children in public or private institutions or other care facilities, are educated with children who are not disabled, and special classes, separating schooling or other removal of children with disabilities from the regular education environment occurs only when the nature or severity of the disability of a child is such that education in regular classes with the use of supplementary aids and services cannot be achieved satisfactorily (p. 24).

It is important to note that IDEIA does not require full inclusion of all children in the regular classroom, nor does the term *inclusion* appear in the writing of the law; however, the legislation does require school administrators to consider modifications in the regular classroom before moving the child to a more restrictive placement. IDEIA requires an individualized placement decision for each student and does not support one-size fits all approaches to placement.

As the population of students in special education in the public schools continues to grow and school district administrators continue to work to develop special education programs in compliance with IDEIA, so do the legal disputes and case law decisions that continue to shape the field. In fact, the majority of court decisions that reference special education placement in the LRE side with the student with disabilities. Interestingly, little consideration is given to how

placement of this student in the inclusive setting could influence the academic performance of the non-disabled students in the classroom.

In 1989, *Daniel R.R. v. State Board of Education 874 F. 2d 1036 (5th Cir.)*, Daniel, an elementary student with Down Syndrome, was fighting the school district's decision to remove him from an inclusive setting because they claimed "it was of no benefit to him because of his academic performance" (p. 18). Results of this case rendered a two-part decision process to determine the appropriateness of placement.

First, a school district must ask the following questions:

1. Has the school taken steps to provide supplementary aids and services to modify the regular education program to suit the needs of the disabled child?

2. Once modifications are made, can the child receive an educational benefit from regular education?

3. Will any detriment to the child result from placement in the regular classroom?

4.What effect will the disabled child's presence have on the regular classroom environment and thus on the education the other students are receiving?

Second, if the decision is made to remove this child from the classroom, then the following question must be asked:

5. Has the child been mainstreamed to the maximum extent possible?

As a result, the courts sided with the parents, claiming that the school district is not acknowledging the non-academic benefits of an "inclusive setting."

Years later, during the case of *Sacramento City Unified School District, Board of Education v. Rachel Holland. 14 F3.d 1398 (9th Circuit. 1994)*, the 9th Circuit Court used four factors to determine appropriate placement:

1. The educational benefits available to the child in the regular classroom
2. The nonacademic benefits of interaction with children who are not disabled
3. The effect of the disabled child's presence on the teacher and other children in the classroom
4. The cost of mainstreaming

After careful consideration, judges explained that caution must be taken when looking at Criteria 3 and 4 because if wrongly interpreted, then the legal mandates of IDEA could be violated. Additionally, important doctrine resulted from the decision of *Sacramento v. Rachel Holland (1994)*: "The effect of the presence of the child with disabilities on the other children in the classroom should be a concern only if the child is so disruptive or requires so much of the teacher's time that the teacher is unable to teach" (Martin, et al., 1996, p. 39).

Despite the courts criteria and the mandate in IDEA to offer a "continuum of alternative placements," a decision from a case two years later afforded school district officials to have the power to remove classified students from inclusive placement. *Clyde K. v. Puyallup School District, 35 F3d 1396 (9th Circuit, 1994)* rendered the decision that "while school officials have a statutory duty to ensure that disabled students receive an appropriate education, they are not required to sit on their hands when a disabled student's behavioral problems prevent both him and those around him from learning" (p. 34). In other words, school administrators have the right to act when a child who is in a learning-restrictive environment (LRE) is influencing the learning of the other students in the classroom. This decision later coincided with IDEA's

regulations: "If the child's behavior in the regular classroom, even with the provision of appropriate behavioral supports, strategies, or interventions, would significantly impair the learning of others, that placement would not meet his or her needs and would not be appropriate for that child" (p. 234).

As case law continues to drive changes and legalities with regards to special education and the need for the creation of "inclusive classrooms," it is not the only force driving school reform. In fact, since the release of the document of *A Nation at Risk* in 1983, the federal government has continued to create reports and pass influential legislation shaping the future of public education in the United States.

In August of 1981, President Ronald Reagan formed the National Commission on Excellence in Education. This group was charged with analyzing data from schools across the nation in order to determine the effectiveness of America's schools. The report, which was released in 1983, established schools across the country as "at risk" for failing the youth of America, especially minorities. A summary of the report indicated that 13% of all 17-year-olds in the United States could be considered "functionally illiterate," SAT scores were continually declining, 40% of students were unable to draw conclusions from written text and more students were in need of remedial courses in college" (U.S.D.O.E., 1983, p.76). Findings of the report caused national panic and the call for immediate reform of schools. Members of the U.S. Department of Education promised "the best effort and performance for all students, whether they are gifted or less able, affluent, disadvantaged, whether destined for college, the farm or industry" (U.S.D.O.E., 1983, p. 2). It is important to note that *A Nation at Risk* was the beginning of an era in achievement testing and standards-based reform.

The following decade continued the movement towards standards-based reform with The Improving of America's Schools Act of 1994 (IASA). The IASA reauthorized the Elementary and Secondary Education Act of 1965 (ESEA) and maintained the same focus--low achieving students in poor schools, high standards for all children, resources targeted to areas of greatest needs, and flexibility coupled with responsibility for student performance. Part of the ESEA and later the IASA was Title I, which was a section of the Act focused specifically on the academic achievement of disadvantaged students who were not proficient in basic skills.

In addition to IASA, another landmark legislation, Goals 2000: Educate America Act, was passed in 1994. This act focused not only on the disadvantaged students but on the needs of all students. Writers of the act specifically looked at educational equity for children with special needs. They claimed that "the de facto segregation of students into regular classrooms and special services classrooms had to end" (USDOE 1983, p. 12). The Act mandated accountability for all schools and began an accountability system to identify schools that were not helping all students perform proficiently.

As standards-based reform continued to shift the education paradigm to an essentialist environment focusing on school accountability by means of standardized test outcomes, on January 8, 2002, President G. W. Bush signed into law the most dramatic reauthorization of the ESEA, entitled the No Child Left Behind Act of 2001 (NCLB). The preamble of NCLB outlines the new mission to improve schools: "An act to close the achievement gap with accountability, flexibility, and choice, so that no child is left behind" (p. 15). The NCLB legislation emphasizes accountability based on the following conditions: "challenging academic content and student achievement standards, all students will be tested in mathematics, reading, or language arts and science, adequate yearly progress (AYP) must be demonstrated so that by 2013 all students meet

standards, and separate measures must be reported for all (1) k-12 children, (2) economically disadvantaged, (3) students from major racial and ethnic groups, (4) students with disabilities, (5) students with limited English proficiency" (Orlich, 2004, p.88). With that said, by the year 2014 all students must meet the target of 100% proficiency which, in addition to the general education population, includes each of the mandated subgroups listed above. If school leaders do not meet annual AYP targets on the way to 100% proficiency, the schools are labeled as "a school in need of improvement." If this label continues for five consecutive years, NCLB has strict, costly consequences called "corrective actions."

In summary, driven by case law and federal education legislation such as NCLB and IDEA, principals and school leaders face difficult demands in order to make every child succeed and abide by laws requiring education in a LRE. They continue to strive to make educated and data-driven decisions on how to improve the academic achievement of all students, including children with disabilities in order to make AYP; however, although inclusion and mainstreaming are legally mandated and has some research suggesting positive achievement results for classified students, what influence does inclusion have on the general education population? NCLB requires improved academic achievement and equity for all students.

School Variables

Empirical Research on the Effects of Inclusion on the Academic Achievement of Students without Disabilities

As policymakers continue to create legislation with rigorous accountability measures, school officials struggle to determine the academic influence of inclusion on students without disabilities. In addition, they must determine annually what general education students are placed in inclusive and non-inclusive classrooms and whether or not this placement will

influence their academic achievement. Current practices used by administrators are not researched-based, nor do they have empirical evidence to support them. In fact, there is little research examining the influence of inclusion on the academic achievement of general education students, and the research that does exist is problematic in both design and methodology, being descriptive or quasi-experimental (Peltier, 1997). Interestingly, Staub and Peck (1994) conducted a review of the research on academic outcomes for non-disabled students and noted that although the research is limited, they were encouraged by the consistency with which existing studies indicate that inclusion does not harm non-disabled children (Peltier, 1997).

Research in this area began in 1988 when Affleck, Madge, Adams, and Lowenbraun used a non-equivalent control group design on 13 elementary school classrooms, 3 buildings, in Grade Levels 1-6 within the Issaquah School District in Washington to examine how disabled and non-disabled students achieved academically when exposed to the Integrated Classroom Model (ICM).

The ICM model was designed to educate mildly handicapped children in the same classrooms with regular education children for the entire school day (Affleck et al., 1988, p. 342). In the ICM setting, regular school district curriculum was used, teachers were exposed to both regular and special education settings, and a half-time aide was assigned to each ICM classroom. The integrated classrooms were composed of one third mildly handicapped students. To be determined "mildly handicapped," the students had to meet state eligibility requirements to be deemed learning disabled, mildly mentally retarded, and/or seriously behaviorally disabled. The other two thirds of the classroom were general education students. In other words, the classroom size was approximately 24 students, 8 of whom were "mildly handicapped."

This study examined an experimental group of regular education students and mildly mentally retarded students assigned to an ICM classroom while a contrast group of randomly selected regular education students of the same size and grade level were enrolled in a regular classroom setting without disabled peers. There were 39 regular education students in Grades 3 and 4 from one building and Grade 5 from another. All students in both groups were Caucasian and had equal socioeconomic status and placed in an ICM setting for one year followed by placement in a general education classroom.

Each student within the group was administered the California Achievement Test Battery as a pre-test, post-test measure. Once the percentile scores were measured, they were converted to NCE scores and an analysis of variance was used on the pre-, post-, and gain scores. Researchers found that there was no significant difference between the groups of regular education students in the ICM classrooms and the regular classrooms in reading, language, or mathematics achievement or between those students in an ICM classroom and those in a non-inclusion classroom.

While this study begins initial inquiry in attempting to provide empirical evidence showing that there was no significant difference between the experimental group (general education students enrolled in ICM classrooms) and the control group (general education students not enrolled in ICM classrooms), it does not come without limitations. First, the findings of this study cannot be generalized to urban, rural, or culturally diverse school settings because of the sample of convenience comprised of all Caucasian students with the same SES. Because of the homogeneous composition of the sample, it may not provide typical results if this study was reciprocated at another location. Second, the researchers mention that the school district already has a predetermined philosophy of integrative practices; therefore, this study

could be subject to bias and skewed results based on the predisposition and exposure to this philosophy. Finally, results of a study can be deemed worthwhile based on the fact that they can be replicated. Because this study provides limited details and methods on how ICM is implemented and how the experiment was conducted, it would be difficult to replicate, thus causing some questions concerning the reliability of the results.

Sharpe, York, and Knight (1994) continued to investigate the influence of inclusive school environments on the academic performance of general education students. The researchers rationalized this study by stating that although there is research showing the benefits of inclusion for students with disabilities, there is continuing controversy on whether the inclusive movement has an academic influence on students without disabilities. Within this question is an unanswered inference: "Does accommodating the needs of a few place at risk the learning opportunities of the majority?" (Sharpe et al., 1994, p. 284). As a result of this debate, researchers set out to answer these questions: (1) Will test scores of children without disabilities go down when children with disabilities are included in general education classrooms? and (2) Are classroom teachers more likely to see increased behavioral problems when students are educated in an inclusive environment?

For the purposes of this study, the researchers defined an inclusive environment as "a general education elementary classroom with members including some who had significant disabilities and who were previously taught in self-contained special education classrooms" (Sharpe et al., 1994, p. 286). A post hoc quasi-experimental pre-test, post-test design was used to look at the experimental group of general education students in an inclusive setting and the control group of general education students in a non-inclusive setting.

The participants of this study were enrolled in a K-6 elementary school in Minnesota where class sizes were composed of approximately 30 students. The student population was not diverse, being 96% White and 4% Native American. Additionally, the SES of the population was 80% middle class and 20% designated as below the national poverty line.

There were five students in special education chosen to be a part of this study, each having been previously educated in a self-contained classroom. Three of the five students were defined as having moderate to profound mental retardation, one student was classified as having "an educable mental handicap and had additional challenges with interventions by social service organizations, and the fifth student had a severe emotional disorder (Sharpe et al., 1994, p. 288). These students, previously educated in a self-contained classroom, were now in a general education setting for at least 80% of the school day with their general education peers. Each classroom received one of these students while one classroom had two of the students in special education.

The regular student sample was comprised of 143 general education students who at the time the data was utilized were in Grades 3 or 4. The sample was 49% male and 51% female. Thirty-five general students represented the inclusion group, and 108 students were in the comparison group. The researchers followed specific procedures in determining which data to include and exclude in order to ensure valid results. As a result, six students were removed from the sample due to incomplete data.

To determine results, the researchers used four performance measures: (1) the Science Research Association Assessment Survey (SRA) to compare pre- and post-academic performance of the two groups, (2) reading level determined by assigning ranks to each student's

reading book placement using Houghton Mifflin Company basal readers, (3) academic performance as indicated by grades on report cards for the areas of reading, mathematics, and spelling and (4) general performance as indicated by conduct and effort grades on report cards (1994).

Using a one-way analysis of variance (ANOVA) and Pearson chi-squares, the researchers determined results. Researchers' results yielded the following: the results of the pretest ANOVAs looking at SRA achievement tests showed no significant difference between the control and experimental groups ($p>.05$). Additionally, the results of the chi-square analysis using report card data showed no significant difference between the two groups in the following areas: reading (chi-square 0.041, $p=.84$), spelling (chi-square 3.031, $p=.22$), mathematics (chi-square 0.002, $p=.96$), conduct (chi-square 3.759, $p=.15$), effort (chi-square 0.916, $p=.63$).

From the results, the researchers conclude that there was no indication of a "decline in academic or behavioral performance of classmates educated in inclusive classrooms on the assessment tools used" (Sharpe et al., 1994, p. 286). Also, it is noted that this study is only an initial step in answering the larger question and additional studies must be conducted to look at a range of variables that exist within the general education population.

Although this study adds to the small body of research that exists on this topic, its limitations cause some underlying questions and concerns. First, the sample was very limited in terms of demographics. The students were mostly White, middle class students with a small population of Native American and lower SES children, which does not render the results of a diverse community of learners. Also, researchers did not have a balanced sample of general education students "in" and "not in" inclusion settings. In fact, only 35 students represented the

inclusion group versus the 108 that represented the non-inclusion group. Because of the skewed sample, the results may have been different given a more even distribution of students in the experimental and control groups. Finally, there were only one or two students in special education "included" in each setting. This is not typical in most inclusion settings, which have a higher ratio of students in special education in the classroom. In this study the special education to general education ratio is 1/2:30, whereas in other school districts the ratio may be more like 7:30. These mentioned limitations cause questions in terms of validity of the results and require further clarification and investigation in order for the results to be accepted and generalized.

Daniel and King (1997) attempted to define how "special education interfaces with general education" (p. 73) by following the research of Sharpe, York, and Knight. The researchers investigated the effects of students' placement versus non-placement in an inclusion classroom by 4 dependent variables, including the following: (a) parents' concerns about their children's school program, (b) teacher and parent reported instances of student behavior problems, (c) students' academic performance, and (d) students' self reported self-esteem. In addition, researchers looked at whether student placement in different types of inclusion programs would result in differences in the dependent variables.

Using a quasi-experimental design of Grade 3-5 students (n=207) from interactive rooms (rooms that had been formed according to criteria established by educators at the given schools), which resulted in the elimination of random assignment. The sample was comprised of three groups: Group 1 - n=68 students from four non-inclusion classrooms, Group 2 - n=34 from two clustered inclusion classes and Group 3 - n=105 from six random inclusion classes.

Standardized instrumentation consisted of the Child Behavior Checklist (CBCL), a self-administered 113 item instrument completed by a parent, teacher, or other caregiver looking at the child's "adaptive functioning or problems in a standardized format," (Daniel & King, 1997, p. 72), the Self Esteem Index (SEI), a 30-minute paper and pencil assessment looking for responses to 80 items as *always true, usually true, usually false, or always false,* and lastly the Stanford Achievement Test (SAT), a standardized assessment for students in Grades 7-9 (Daniel & King, 1997, p. 72). Researchers used the total battery scores in mathematics, reading, language, and spelling from spring 1993 and spring 1994, which were prior to or at the end of a year-long inclusion/non-inclusion placement.

Data were collected via the students' archival records, giving students the exam during the regular school hours and teachers completing the Teacher's Report Form (TRF) data on each child and returning it to the researchers. Data were then analyzed using a discriminate analysis. Complete data were only available for 178 out of the 207 students, divided into 63 third graders, 52 fourth graders, and 63 fifth graders.

Results indicated that "(a) parents of students in the inclusion classes expressed a higher degree of concern for their children's school programs, (b) teachers and parents of the students in the inclusion classes reported more instances of behavioral problems, (c) students in inclusion classes were more likely to experience gains in reading scores with no noteworthy differences for mathematics, language, and spelling, and (d) students in inclusion classes reported lower levels of self-esteem" (Daniel & King, 1997, p. 67). The discriminant analysis of data (SEI, SAT, & CBCL – teacher response) indicated: Grade 3 effect size of 34.6% (Wilks's Lambda Λ=.65), $p < .01$, indicating a statistically significant difference in the performance of the non-inclusion and

inclusion students; Grade 4 effect size of 31.2% (Λ=.69), $p < .10$, indicating a statistically significant difference in the performance of the non-inclusion and inclusion students; and Grade 5 effect size of 37% (Λ=.63), $p < .01$, indicating a statistically significant difference in the performance of the inclusion and clustered students (a comparison of two different inclusion models).

The analyses of data resulted in a significant difference in performance of the students in inclusion classes and non-inclusion classes. The researchers reported that the results of the study were mixed and somewhat difficult to decipher as, the Grade 3 inclusion students made gains in reading, experienced more behavioral problems, and reported lower levels of self-esteem when compared to the non-inclusive students; Grade 4 inclusion students made gains in mathematics and reported higher self-esteem scores when compared to the non-inclusive students; and Grade 5 students yielded mixed results. The researchers concluded that no consistent pattern of academic performance emerged, and the higher incidence of behavior problems in the inclusive classrooms may diminish time on instruction as a result of time devoted to handling these problems. Consequently, the behavioral issues brought into the classroom by special needs students may have a negative effect on their classmates.

Although this research continues to add to the limited body of existing research, there is concern regarding the lack of demographics addressed. The researchers fail to mention the makeup of the student population and variables such as socioeconomic status, gender, and ethnicity, which may have an effect on the data and results. Additionally, there is no mention of the structure of an inclusion classroom; e.g., how many teachers are present, classroom size, how many special needs students are included in the classroom, and classification (emotionally

disturbed, Autistic, cognitively impaired, etc.) of students with disabilities who are included in the inclusive setting. It is also important to note that the data collection methods are perplexing and unclear at best because researchers mention that the SAT is a Grade 7-9 assessment, yet the student sample were in Grades 3, 4, and 5. In order to validate these results, more detailed procedures must be defined and clarified because the findings of this study are unclear and mixed.

Saint-Laurent et al. (1998) designed a study to evaluate the influence of an in-class service model called Intervention Program for Students at Educational Risk (PIER) on the achievement of students at risk of school failure. A total of 606 White, French speaking, third grade students from 26 schools participated in this study. Each school, located in one of two main urban areas of Quebec, had one general education class selected with an average of 24 students per class. It is also important to note that 276 students were of high SES schools, 148 students were from middle SES schools, and 182 students were from low SES schools.

Researchers used four criteria to determine at-risk students in the treatment and comparison classes: (1) low results on the Grade 3 academic tests of reading, writing, and mathematics, (2) teacher ratings of abilities in reading, writing, and mathematics, (3) grade retention, and (4) identification as students in special education by district and school criteria. The treatment group was composed of 288 students, approximately 145 girls and 143 boys, 79 at-risk, and 34 students in special education (27 learning disabled, 5 behavioral disorders, and 2 hearing impaired). In the comparison group there were 318 students, 139 girls and 179 boys, 86 students considered at risk, and 38 students in special education (32 with learning disabilities, 4 with behavioral disorders, and 2 with communication disorders).

During the 1993-1994 school year, students in the treatment condition, were instructed using only the PIER model, which included four components: (a) collaborative consultation, (b) cooperative teaching, (c) parent involvement, and (d) strategic and adapted instruction in reading, writing, and mathematics with both a general and special education teacher. In contrast, students in the comparative group (non-treatment) were taught by both teachers using regular education teaching methods only. Data were collected for achievement variables in September (pre-test) and June (post-test). Tests sanctioned by the provincial Department of Education were given in reading, writing, and mathematics.

To evaluate the effectiveness of the PIER program, the researchers conducted two multivariate analyses of covariance (MANCOVAs). "A 2 (Group) x 2 (Student Type) MANCOVA performed on these scores revealed a significant effect for Group x Student Type interaction, [F (3, 511) = 7.03, p [is less than] .001, Wilks's lambda = .96]; Univariate tests showed significant differences in reading, [F (1, 513) = 4.96, p [is less than] .05], writing, [F (1, 513) = 5.56, p [is less than] .05], and mathematics [F (1, 513) = 7.24, p [is less than] .01]" (p. 248). Effect sizes were reported as follows: reading (-0.04), writing (0.50), and mathematics (0.13). The effect size for writing is considered moderate; however, the effect size for mathematics is considered low and should be taken into account when interpreting the results of this study.

These statistical findings indicated the following results of the study: (1) the treatment program benefits both at-risk and general education students in at least one academic area, (2) the PIER model was at least as effective as activities conducted in the comparison classes for reading and mathematics and produced higher writing scores for at-risk students, (3) general

education students were not held back by the presence of at-risk students who were present in the classroom, and (4) general education students benefited from the additional interventions that form part of the PIER model in reading and mathematics.

Although these findings give additional insight into the effect of inclusive settings on general education students, there are "some methodological weaknesses" inherent in this study. First, because the study was conducted in real classrooms, the researchers were unable to account for and control all variables present in instruction and in each student. Also, random assignment is essential to the generalizability of these findings and only the teaching dyads (pairs of general and special education teachers) were randomly assigned; therefore, students may have been placed in certain classroom structures or with certain teachers based on pre-determined factors. This lack of random sampling jeopardizes the validity of the data. It is also important to mention that teacher participation in this study was on a voluntary basis; moreover, the teachers selected may not be representative of the larger teacher population. Along with threats to validity, the researchers fail to know the psychometric properties of the achievement tests. Even though these tests are used as a standardized measurement tool across the province, it presents cause for concern. Finally, because this study is evaluative in nature, there are many unanswered questions, such as what is "quality teaching" and what are "effective in-class support services." With that said, this study provides preliminary insight into the effect of inclusive and non-inclusive settings on student achievement, but it must be noted that said limitations may limit the usability of these findings.

Huber, Rosenfeld, and Fiorello (2001) continued to add to the existing body of literature by examining the "differential influence of inclusion and inclusive practices on high, average,

and low achieving general education students" (p. 501). Over three years, a random sample of 477 general education students in Grades 1-5 was either placed in the experimental group (educated in a classroom where students with disabilities were present) or the control group (educated in a classroom with all general education students). The sample of students was primarily working class and 72% White, 27% African American, and 1% Asian. Additionally, 50% of the students were considered eligible for free and reduced lunch. The students with disabilities were considered able to meet the demands of the class with minimal supports. In other words, 41 out of 49 students were considered learning disabled, while the other 9 fell into one of the following categories: behaviorally/emotionally handicapped, educable mentally handicapped, and other health impaired.

Using the scores of the Metropolitan Achievement Test, the researchers made two comparisons. First, they looked at how incremental changes in general education students' achievement scores for high, middle, and low achievers differed between the experimental and control groups. For data analysis, students were grouped into three groups for math and reading based on previous test scores: high-achieving, on grade level, and low-achieving students. A 2 (year) x 3 (skill level) ANOVA, a balanced factorial design with repeated measures on the year factor was performed and Tukey's test was employed to identify significant differences between group means. Huber et al. (2001) reported that inclusive school practices were found to have a differential effect in that low-achieving general education students appeared to benefit academically, while higher-achieving students were adversely influenced.

Results indicated the following: (1) the student skill factor had a statistically significant effect on incremental change in general education students' reading scores, $F(2,498) = 12.86$, p

$< .001$, (2) the student skill factor had a statistically significant effect on math change scores, $F(2,546) = 26.85, p < .001$, (3) no significant differences among group means for general education students in classrooms with different numbers of included students with disabilities, $F(7,791) = 0.87, p = .53$, and (4) significant differences among the group means of children educated with different numbers of included students with disabilities, $F(7,794) = 4.82, p < .001$. The analysis of the data suggests that inclusion and inclusive practices may lead to different rates of achievement for general education students.

From this data, Huber et al. (2001) reported that inclusive school practices were found to have a differential effect; low-achieving general education students appeared to benefit academically, while higher-achieving students were adversely influenced. General education students enrolled in the inclusion classes were not significantly affected in reading; however, the effect was mixed in math. Huber et al. also state that although the results indicated some discrepancy in student achievement between groups in math, the reasoning or cause is unclear.

Huber et al. (2001) used a large random sample with a design including both an experimental and control group. This methodology provided for a strong study with promising results; however, there are a few limitations that could skew the data set and results. First, the majority of students in special education used in the sample were predominantly considered learning disabled. There was a very small portion of students with a variety of other disabilities. A larger, diverse special education population needed to be used in order to be more representative of students in special education nationwide because in some cases general education students did well with larger numbers of included students, while other general education students did poorly (2001). This may have to do with individual differences and needs

within the special education population. Also, Huber et al. disclose that halfway through the study, the reading program was changed to a "whole language approach," which may have caused the reading scores to decrease; therefore, it is difficult to measure whether the drop in reading scores is due to curriculum changes or the inclusive setting.

McDonnell, Thorson, Disher, Mathot-Buckner, Mendel and Ray (2003) conducted an exploratory study using a quasi-experimental pre-test post-test design to examine the influence of "inclusive educational programs on the achievement of students with developmental disabilities and their peers without disabilities" (p. 226). The participants in this study included 14 students, ages 6-12, with developmental disabilities in inclusive classrooms (I.Q. ranged from "not testable" to 78, with an average of 54.6 based on standard I.Q. tests), their 324 typical classmates, and 221 typical students in non-inclusive elementary classes enrolled in five different elementary schools, located in four different districts. The districts were a combination of rural, suburban and urban. Students were enrolled in Grades 1-5, with one to seven students in each inclusion class having a disability.

The researchers measured student performance in two ways: the Utah Core Assessment (UCA), a criterion referenced achievement test, and the Scales of Independent Behavior – Revised (SIB-R) (used with developmentally delayed students). Data analysis was conducted as follows: Pre-test and post-test performance changes on the SIB-R were assessed using a Wilcoxon Signed Rank Test and non-disabled students were compared in inclusive and non-inclusive classes on the UCA using a One-Way ANOVA.

The results of a one-way analysis of variance indicated no significant difference in academic performance for non-disabled students enrolled in inclusion classes and their non-

disabled peers enrolled in non-inclusive classes (reading/language arts, F=.02, p=.87, df = 1,543; mathematics F=.39, p=.52, df = 1,543). "The results suggest that the presence of students with developmental disabilities did not negatively influence the educational achievement of students without disabilities" (McDonnell, 2003, p. 235). Furthermore, the results of the two-tailed Wilcoxon Signed Rank Test on pre- and post-test measures were found to be statistically significant (Z=3.18, $p < .001$), indicating that students with developmental disabilities made gains in adaptive behavior. Researchers concluded that their results were consistent with the results found by Sharpe et al. (1994), suggesting that "the concerns that some authors have expressed about the negative influence of inclusion on educational achievement of students without disabilities may be unwarranted" (p. 235).

Although the results of this study provide additional empirical evidence that the academic achievement of general education students in inclusive settings is not negatively influenced, it is important to note the limitations of this study. First, there are a small number of students with developmental disabilities included in this study; therefore, the generalizability of these findings may not be representative of the larger special education population. Also, the faculty participating in this study had motivation to expand inclusive education. They were provided training and technical assistance during the course of the study. It is possible that results may be different in a school where the staff is resistant to these practices.

Fletcher (2010) aimed to examine the spillover effects of inclusion on non-disabled classmates. He indicates that although inclusion has been a federal policy initiative, there is little research examining the effects of this policy on non-disabled classmates during early elementary grades. The Early Childhood Longitudinal Study, Kindergarten Cohort (ECLS-K) is a nationally representative sample of kindergarteners, their teachers and schools. The sample consisted of

(n=11,373) students from both public and private schools who attended both full-day and half-day programs and were from diverse socioeconomic, racial/ethnic backgrounds. Additionally, over 25% of the sample had a classmate with a learning disability, and 10% had a classmate with an emotional disorder. This study controlled for student-level fixed effects and in so doing identified spillover effects, using student differences in exposure to classmates with emotional problems between kindergarten and first grade.

An ordinary least squares regression was completed to examine the relationship between mathematics and reading tests and the effect of having a classmate with a serious emotional problem. Cross-sectional results indicate that having a classmate with an emotional problem reduces reading and math scores at the end of kindergarten and first grade by over 10% of a standard deviation, which is reported as one-third to one-half of the minority test score gap. "The magnitude of this effect is approximately 40% of the adjusted Hispanic–White average difference in test scores and approximately 25% of the size of the adjusted Black–White test score gap" (Fletcher, 2010, p. 77). The researchers further reported that for mathematics scores, Black and Hispanic students seem to be most affected by exposure to classmates with emotional problems (12% and 9% of a standard deviation, respectively). For reading achievement, White and Black students were reported to experience similar decreases with exposure to classmates with emotional problems (3% of a standard deviation); however, the effects on Hispanic students were nearly 10% of a standard deviation. The results were also very comparable by gender, education level of mothers, and marital status of mothers. Finally, females were more affected than males (4% vs. 2%), and students with more highly educated mothers were also more affected (4.5% vs. 1%). Additionally, the results for male math regression scores = -0.066 at 1%

and female math regression scores = -0.053 at 5%; male reading regression scores = -0.013 and female math regression scores = -0.043. The results suggest that in both reading and mathematics, exposure to girls in the classroom increases achievement; a 10% increase in the proportion of classmates who are girls increases achievement by 1% of a standard deviation.

According to Fletcher (2010), the consistent result for mathematics and reading test scores indicates that students with classmates who have a serious emotional problem score significantly lower than other students. The author contends that the results suggest that the policy of full inclusion of students with all types of disabilities into the regular education classroom may need to be re-evaluated. In particular, the benefits and costs both to the disabled children and the non-disabled children should be considered.

Although Fletcher (2010) used a large and diverse sample, limitations were noted. "The study was limited by the fact that no information was provided in the data describing the algorithms used by schools to match students with classmates and purposeful matching of students could introduce bias in estimating spillover effects" (p. 81). In addition, this study was limited by the inability to control for teacher sorting across classrooms within schools.

Trabucco (2011) conducted an Independent Samples t test to examine to what extent placement in a co-taught inclusive setting correlate with non-disabled students' academic achievement. The participants in the study were enrolled in grade 3 and from an upper middle class suburban elementary school district in New Jersey. The sample population in the inclusion classroom included the following demographics: 0 students in the inclusion setting qualified for free and reduced lunch while 1 student was eligible in the non-inclusion classroom, 5 students were classified as special education in the inclusion classroom, and 4.9% of the students in the inclusion setting were eligible for basic skills versus 6.1% in the non-inclusion setting. Trabucco

also makes note of the pre-achievement mean scores of 30.33 in the inclusion setting and 31.02 in the non-inclusive setting.

An analysis of the influence of the independent/grouping variable (placement in an inclusion setting) on the dependent variable (academic performance on the New Jersey Assessment of Skills and Knowledge) was conducted specifically in the academic area of mathematics. Trabucco examined the measures of central tendency (the average performance of each placement group on the NJ ASK mathematics scores) and the standard deviation of those scores. Additionally, t tests were conducted to look for significant differences between the two comparison groups.

Results found that there was no statistically significant difference between overall achievement scores on the NJ ASK mathematics test for students placed in inclusion and non-inclusion classes (t = .612 with 97 *df* and p = .542). The effect size was 0.17. Trabucco also examined a subset of questions looking at the various cluster of questions asked on the NJ ASK including the following: Number and Numerical Operations, Geometry and Measurement, Patterns and Algebra, Data Analysis and Discrete Mathematics. Data found that there was a statistically significant difference between Number and Numeric Operations cluster scores on the NJ ASK for students placed in inclusion and non-inclusion classes (t =1.941 with 30.2 *df* and p = .042) with a relatively small effect size of 0.37; however, the rest of the clusters (Geometry, Algebra, and Discrete Mathematics) were not found to be statistically significant.

Overall, Trabucco concluded that placement in a co-taught inclusive classroom does not influence the achievement of non-disabled students in mathematics with the exception of performance on Number and Numeric Operations when prior [pre] achievement is controlled. One possible explanation for the results is that "general education teachers and special education

teachers bring a tremendous amount of knowledge and skills to the task of teaching, and by being paired together, they pool their expertise" (Luzader, 1995, p. 19). It is also important to note that the study is consistent with other research that has been conducted in this area.

Although this study continues to add to the limited body of research on the influence of the inclusive setting on student achievement, there are limitations which must be noted. First, there is a lack of random sampling which limits the researcher's ability to make inferences about the performance of the larger group [population]. Second, participants lacked diversity in terms of cultural background and socioeconomic status and therefore the findings cannot be generalized to urban, rural, or culturally diverse school settings. Also, the researcher makes little note of the types of disabilities possessed by the students in special education. Previous studies have noted that certain classifications of disabilities are known to affect classes differently. Third, the findings of this study remain tentative because of the small sample size of the inclusion group (n=15).

In summary, The Individuals with Disabilities Education Act (IDEA) requires that emphasis be placed on instructing students with disabilities within general education settings to the maximum extent appropriate [34 CFR sect. 300.500 (b)(1)]. Additionally, court rulings have also supported more inclusive placements. Although recent research has noted the academic, social, and emotional advantages of inclusion for students who are deemed "disabled," Fletcher (2010) states, "While there is mixed evidence on the effects of inclusion policies on the students with special needs, research examining potential spillovers of inclusion on non-disabled classmates has been scarce (p. 69)." In this instance, scarcity has become complexity due to the mixed evidence presented by researchers.

The body of empirical research that does exist concerning the influence of inclusion on non-disabled students' academic achievement has resulted in mixed outcomes (Daniel & King, 1997; Huber, Rosenfeld, & Fiorello, 2001; Hunt, Staub, Alwell, & Goetz, 1994; Kalambouka, et al., 2007; Manset & Semmel, 1997; Saint-Laurent et al., 2002; Sharpe, York & Knight, 1994; Staub & Peck, 1995). Some researchers claim that inclusion does not influence the academic achievement of non-disabled students; others find that non-disabled students are negatively affected, while others argue that inclusion improves non-disabled academic achievement. The studies and results vary by the population being served (type and level of disability), content areas studied (math, language arts), and other variables such as model of inclusion, years of teacher experience, class size, ethnicity, pre-achievement, and sample size. As concluded by all researchers, additional study is needed in this area to continue to add to the scarce body of research examining how inclusion influences the academic achievement of non-disabled students. As stated by Huber (2001), "With a better understanding of how specific policies affect classrooms, school personnel could use resources to serve students more equitably and efficiently" (p. 502). This study aims to clarify and add to the current body of research on how an inclusive setting influences the academic achievement of non-disabled students.

Classroom Peer Effects on Student Achievement

Before the U.S. Congress passed an amendment to the Vocational Rehabilitation Act in 1973 that included a provision prohibiting discrimination against persons with disabilities in local and federally assisted programs and activities, students with disabilities were educated in "institutions" or "homes" assigned on the basis of the severity of their disability. In other words, students with disabilities were separated from their non-disabled peers and forced into

homogeneous groups consisting of others with the either the same disability or worse. As legislation such as IDEA and ADA developed, "inclusion" continued to become present throughout the public schools across the country, allowing students with disabilities a free and appropriate education (FAPE) in their local neighborhood school.

As inclusion became mandated by legislation, school officials emerged as unsure of how to handle "including" students with disabilities in their schools. Because the common thought was that having mixed-ability grouping in classrooms would lower expectations and standards for the non-disabled students, school leaders and administrators attempted to segregate individuals with disabilities into self-contained or resource room settings; however, research has shown that "every means of grouping students by ability or performance level has drawbacks that may be serious enough to offset any advantages" (Slavin, 2008, p. 458). Teachers to administrators have typically tried to reduce variability by assigning students to classes based on some indicator, whether it be gender, ability level, or a variety of test scores (Zaharias, Achillies, Cain, 1995). Unfortunately, this practice is the opposite intention of inclusive practices. In other words, even though children with disabilities were "included" in their neighborhood schools, they were not exposed to the same education, social surroundings, and benefits of their non-disabled peers. Research has demonstrated over time that ability grouping may stigmatize low achievers, put them in classes or groups for which teachers have low expectations or lead to the creation of academic elites (Slavin, 1988). As a result, grouping may doom children who are not in top tracks to second-class instruction and ultimately deprive students of the examples and stimulation provided by heterogeneous classes (Slavin, 1988; Zaharias, Achillies, & Cain, 1995).

While parents of students with disabilities continued to fight for equality, parents of general education students refuted arguments with the thought that including students with disabilities would change or affect the learning and achievement of their children. Thus, with the practice of inclusion eminent, researchers began to study the effects of homogenous and heterogeneous groupings on academic achievement as the current interest on multi-cultural and inclusive education continues and the debate between homogeneous versus heterogeneous groupings has again emerged as an important consideration for educators and leaders (Zaharias, Achillies, Cain, 1995).

Curious about the effects of middle school ability grouping on the cognitive achievement of students in mathematics and science, Hoffer (1992) conducted a study comparing grouped and non-grouped schools. Hoffer's research is an integral part of the empirical evidence in this area because prior to 1992, most studies (with the exception of about 10, as evidenced by the work of Mosteller et. al., 1996) did not contain a control group; therefore, the quality of the research methodology and the validity of the results were jeopardized. With that said, Hoffer examined two important areas: (1) whether ability grouping raises the aggregate level of student achievement and (2) whether ability grouping increases the learning of all students or jeopardizes the learning of specific groups of students.

Data for this study were collected by the Longitudinal Study of American Youth (LSAY) from fall 1987 to fall 1989. The base sample began with 3,116 seventh grade students and 2,829 tenth grade students drawn from 51 pairs of middle and senior high schools (1992). Also, a survey was distributed to teachers asking for professional background experience and characteristics of their schools. Because of the length of the research, 712 from the original

sample were excluded because they moved or dropped out of school, and 218 students had quit the study before it was finished. Also, demographic variables were not provided because the author controlled for the following, which were collected by survey: family SES, parental education, student gender, and race/ethnicity.

In order to research the hypothesis that students in ability grouped schools learn more on average than students in non-grouped schools and, compared with similar students in non-grouped schools, students placed in low, middle, and high groups will realize advantages or at least no losses in cognitive growth, Hoffer used the NAEP derived LSAY tests designed to measure achievement in both math and science (1992). It is also important that about 40% of the schools used ability grouping for science instruction while 80% used ability grouping for mathematics instruction in seventh grade. These rates continue to climb because in eighth grade about 50% of schools use grouping for science, while 92% use grouping for math.

In examination of the research questions, Hoffer found that if ability grouping works, then the results should be positive; however, his data show negative effects (Model 1, -0.535 and Model 2, -.0006) and that none of the estimates approach minimal standards for statistical significance. Additionally, these data show that overall grouping does not work; however, there may be advantages for specific groups of students. Data for the alternative theory known as "differential benefits" do show that grouping has advantages for students placed in higher groups, but does not benefit lower groups of students in both science and mathematics (Science: $p<.05$, $b=2.263$, Math: $p<.05$, $b=2.627$).

Although the author used a design that contained a control and treatment group, it is important to note the limitations involved with this study so that fellow consumers can use care

and caution when applying results. First, the sample of schools is relatively small and may not be a good representation of the American middle school population. Many more schools use homogeneous grouping in middle school; therefore, it is at times difficult to find comparison groups who do not use ability grouping models and schedules. Second, the author discloses that he may not have adequately controlled for initial differences among grouped and non-grouped students as well as controlling accurately for school differences. Additionally, the author notes that he did not look into other ways that school leaders could potentially group students throughout the day such as within class grouping. These factors could skew the data and jeopardize results.

In summary, this study provides evidence that "ability grouping has shown no significant overall benefit in either science or mathematics"; however, differential effects are found in both subjects, though the results in science are less strong (1992). Results hinted toward evidence that in both subjects students in the "high" group learned somewhat more than their peers in "lower groups." Although ability grouping continues to be used by teachers and administrators to cope with a diverse set of students, this study shows that "tracking" can hurt the lower students and may slightly enhance learning for the "high level students." "Overall, it can be said that ability grouping in seventh and eighth grade math and science is not an optimal arrangement when compared with the non-grouped alternative.

In 1996, Mosteller, Light, and Sachs published an extensive literature study in the *Harvard Educational Review* "exploring the nature of the empirical evidence that can inform school leaders' key decisions about how to organize students within schools" (p. 816). Mosteller et al. intended to examine evidence from a variety of studies that supported "skill grouping"

since it is a widely used practice in schools across the country; however, to their dismay, the authors found only a handful of well-designed studies exploring the academic benefits of tracking and of these the results were mixed (1996). Results indicated that the evidence on the influence of grouping students by skill level is limited and much more intensive research must be done in order to make valuable conclusions which influence sound educational practices.

The review began with a look into the various types of grouping occurring in public schools today. It is important to note that although this review focused on 4 forms of skill grouping, for the purposes of this study only three of the four will be discussed given the notion of "inclusive practices."

Whole-Class or Mixed Grouping, Heterogeneous Grouping within Grades

This practice has all students in a grade taught in a group. If the grade level includes too many children for one classroom, the children are split into multiple classrooms representing the whole spectrum of students' skills. This type of grouping produces heterogeneous classes because of the varying skill levels and because the structure is commonly whole-class instruction. Commonly will be used as a control group for a variety of studies.

Between-Class Grouping or XYZ Skill Grouping, Homogeneous Grouping within Grades

In this form of grouping, students in one grade level are divided into three groups of skills: high, medium, and low. The determination into a skill group is made either by pre-achievement or a form of standardized assessment as a whole or by content area. In this case, there is a set curriculum for each group and in some school districts, specific courses are constructed for gifted students or students with special needs.

Within-Class Grouping or Homogeneous Grouping within Classes

In this last form of grouping, the teacher has a class composed of heterogeneous students which he or she sorts into homogeneous sub-groups based on their skill set in the classroom. The difference between "within-class grouping" and "XYZ grouping" is that the three groups stay within the regular classroom for instruction and the teacher may choose to teach a different skill to each subgroup depending on each group's needs. Once the lesson is completed, the teacher may bring the class back together and discuss the day's work; students may have different assignments, but the overall objective is identical.

Mosteller, Light, and Sachs (1996) used specific criteria in order to choose quality research. First, each study had to be an actual experiment that compared learning in skill-grouped classes with learning from whole-class groupings in a school or several schools and had to have a treatment and a control group. Second, the study had to be designed as a randomized field trial; in other words, the assignment of the treatment had to be randomized or be a close approximation to randomization (1996). With these criteria stated, Mostellar, Light, and Sachs made note of a variety of studies that had to be excluded because of the studies having no comparison group, nor a "matched design methodology." In all, 10 studies were discovered that met the above criteria comparing the effectiveness of XYZ grouping with that of whole-class instruction that were carried out between 1960 and 1975. Needless to say, the authors were discouraged by the small body of quality research existing in this area and that the majority were modest in terms of size and scope.

In examining the findings of the 10 studies in which students are grouped, the authors computed effect sizes and found mixed results. It was determined that five of the studies favored skill grouping, three favored whole-class grouping and two give effect sizes near zero.

Additionally, the authors wanted to understand how XYZ grouping affects high, medium, and low skilled students. Again, the authors determined effect sizes for each study and it was found that there is a slight tilt towards skill grouping being more favorable for high skilled students than for medium and low skilled students with effect sizes at 0.08 for high, -0.04 for medium, and -0.06 for low skill groups. Because these are estimated effect sizes, the differences are not to be taken as a firm research conclusion; instead, it should be seen as there is a possibility that skill grouping is favorable for high skilled students (See Table 2).

In summary, results from the 10 chosen studies suggest that "XYZ grouping seems modestly preferable to whole-class grouping for high skilled students, while medium and low skilled students may learn a little more from whole class instruction than with skill grouping" (Mosteller et al., 1996, p. 817). Findings of this analysis support that skill grouping benefits only highly skilled students; however, this must be viewed with caution because of the variability of the findings and the limited number of studies conducted that meet this review's criteria for inclusion. It is also important to note that, on average, XYZ grouping does not have much effect on student achievement.

Mostellar et al. (1996) state that the "main contribution of their examination of the literature of skill grouping is a sharpened awareness of the limited amount of rigorous investigation that has been done and that the amount of evidence for or against skill grouping before the 1990s is scarce" (p. 822). There is not one large-scale, well-designed study that investigates the effects of XYZ grouping over an extended period of time, nor does it follow students for even one year in length. It is essential that more quality studies be conducted so that

results can be clarified and allow school officials to make informed decisions about student placement.

Table 2

Average Performance of Skill Grouped Students

TABLE 2

Average Performance of Skill-Grouped Students as Compared with Whole-Class Grouped Students in the 10 Experiments

Study	*Effect Size**	*No. of Students*
Barton (1964)	.11	204
Bicak (1962)	-.33	75
Drews (1963)	-.04	432
Fick (1962)	.02	162
Ford (1974)	.29	82
Lovell (1960)	.14**	500
Marascuilo & McSweeney (1972)	-.16**	603
Peterson (1966)	-.10	317
Vakos (1969)	.09	184
Wardrop et al. (1967)	.28	82
Sample size weighted average	.00	2641(total)

*Positive effect size favors skill grouping, negative favors whole-class instruction.
** See Appendix 2.

Following the work of Mostellar, Light, and Sachs (1996), who concluded that additional research needed to be conducted in order to create more evidence supporting heterogeneous or homogeneous grouping, Zaharias, Achillies, and Cain (1995) conducted a study using the data base from the Student Teacher Achievement Ratio (STAR) project examining whether random or non-random assignment to classes provides achievement benefits to students in Grades 1-3. Authors of this study expressed concerns that grouping strategies often catered to academically talented students and academically neglected low achieving students. Zaharias et al. stated that "students who are labeled as the 'high' achievers or the 'bright' group tend to be exposed to lower class sizes, more successful teachers, higher expectations, and a more enriched

curriculum" (p. 7). On the other hand, the "low achievers or slow students all too often experience just the opposite" (p. 7).

Although research does not support ability grouping (Mostellar et al., 1996; Slavin, 1990; Zaharias et al., 1995), it is still a commonly used practice throughout many schools across the country. This study aims to contribute to previous research on the ability-grouped class assignment and its effect on student achievement by using a sub-sample of students from school districts that contained both project schools and comparison schools. Zaharias et al. isolated the random assignment variable in order to examine whether random assignment or ability grouping had a positive influence on student achievement.

The STAR database produced a total of 1,157 students available for analysis; n=499 were in random assignments, and n=658 were in non-random classroom placements in Grades 1, 2, and 3. Outcomes were measured by the Stanford Achievement Test (SAT), a norm-referenced test (NRT) in both reading and mathematics as well as the Basic Skills First (BSF), a criterion-referenced test (CRT) developed by the Tennessee Department of Education. Using this data, the researchers used a one-way ANCOVA, controlling for previous test scores, ethnicity, and gender to look for statistically significant effects ($p<.05$).

On examination of the data, Zaharias et al. found the following results: the randomly assigned students outscored the non-randomly assigned students on both tests at each grade level. The most impressive results were the differences in Grade 3, where there was statistical significance on both the SAT ($p<.05$) and BSF ($p<.01$) tests. Additionally, although the math outcomes did not meet the criteria for statistical significance, results favored the non-random group at Grade 1 on both the SAT and BSF. Then, as students entered Grade 2, the experimental

group scored higher on both tests and by the end of Grade 3 the differences were statistically significant in favor of the randomly assigned students (SAT: $p<.001$ and BSF $p<.01$). In total, out of the 18 analyses, 15 favored the randomly assigned students, but only 7 were significant and 3 favored non-random groups, but none were significant.

"Random assignment to classes appears to increase the reading and mathematics achievement of early elementary education students in Grades 1-3. By Grade 3 most of the scores of the random group were higher than the control group scores, especially in math" (Zaharias et al., 1995, p. 12). It is important to note that this study did contain limitations that were openly disclosed by the authors. First, the result of the reading achievement may have been skewed due to an inability to control for homogeneous grouping within a heterogeneous class. Teachers will commonly split students within the same classroom into various subgroups based on reading or language ability; therefore, the results could have reflected some inconsistencies. Regardless of the findings of this study, authors surmise that school administrators and teachers alike continue to use homogeneous grouping as a means of both student placement in classrooms as well as teaching academics to students even though it contradicts empirical evidence acquired.

Although the work of past researchers provided evidence against homogeneous grouping, the practice continues to occur not only in elementary schools, but in middle schools as well. Burris, Heubert, and Levin (2006), concerned with the practice of homogeneous grouping and its influence on student achievement, conducted a longitudinal study examining the effects of providing an accelerated mathematics curriculum in heterogeneously grouped middle school classes in a diverse suburban school district. The authors found that students were being

"tracked" into specific ability-grouped classes and that the accelerated curriculum was reserved for "gifted and talented learners." In contrast, Burris et al. (2006) found research suggesting that school officials should provide a rigorous mathematics curriculum to all students, not only initial high achievers because analyses of international studies such as SIM (Second International Mathematics Study) show that a traditional low-track, remedial curriculum actually depresses the mathematics performance of American students rather than improving it. With that being said, the authors of this study aimed to fill important gaps in the present literature; e.g., would more students take and pass courses at the level of trigonometry and beyond if they took accelerated algebra in eighth grade and would the performance of initial higher achievers decrease if all students were heterogeneously grouped in accelerated mathematics?

Researchers conducted the study in a suburban community in Nassau County, Long Island, where the student population was about 3,500. The population was mostly White, 8% African American, 12% Latino, and 2% Asian with approximately 145 high school students on free or reduced lunch. Of these 145 students that qualified for free and reduced lunch, 98% were students of color. Using a quasi-experimental design, the researchers examined the mathematics achievement data of students in six consecutive annual cohorts. It is important to note that the authors only gave the "treatment" to three of the six cohorts, ensuring for a control group to be used for comparison. The data were retrieved from four different data sources, the ITBS Mathematics Concepts Subtest, scores from the Sequential Mathematics I Regents Exam, students' scores on advanced placement calculus exams, and mathematics courses taken and passed by students in high school.

Examination of the first research question showed that the percentage of students taking advanced math courses did increase after all of the students were exposed to the rigorous math curriculum (2006). In turn, by the end of twelfth grade, 92% of all students in the post universal acceleration group had passed a course and the Regents Exam in high-level math. Additionally, Burris et al. examined the influence of the new curriculum on subgroups of students and found that their findings were crucial for school officials who need to make curriculum-based decisions. After the "treatment" was given to the groups, the number of minority students who met the mathematics graduation requirement tripled from 23% to 75% and while there still remains a gap between the academic achievement of White/Asian students and African American/Latino students, this study showed evidence of this gap narrowing from 46% passing the exam to 67% passing the exam. In summary, the data indicated the following: (1) for all three levels of mathematics courses, membership in a post-universal acceleration cohort was a contributor to the probability of a student taking an advanced mathematics course and (2) once students in middle school had been de-tracked and studied algebra in Grade 8, the probability of a student completing an advanced mathematics course before graduating high school significantly increased (Burris, et al., 2006).

Burris et al. also used a two-tailed t test of means in order to determine whether there was a significant difference between the academic scores of high-achieving students in the control and treatment group. A common concern when using heterogeneous grouping is the influence that it could have on the academic achievement of the gifted and talented students. Analysis of the data for the second research question rendered the following results: the difference in mean scores was found to be not significant at the $p<.05$ level: 93.07 (control group), 91.72 (treatment

group). In other words, the scores of the top-performing students (within the nation's top 20%) were not influenced by heterogeneous grouping. It is important to note that the high-scoring students in the treatment group studied in untracked middle school mathematics classes.

Findings from this study contribute to the limited body of research available on the effects of heterogeneous grouping on the academic achievement of students by providing evidence that de-tracking and exposure to a more rigorous, high-track curriculum is beneficial for all students. It also clarifies some important questions that were left unresolved by Slavin (1990), Hoffer (1992), and Mostellar, Light, and Slavin (1996), claiming that homogeneous grouping provided a slight advantage for higher achieving students. Authors of this study found the opposite to be true; de-tracking had no significant influence on the scores of high scoring students.

Although the researchers used a sound methodology in providing results from both a treatment and control group, it is important to mention the limitations present. First, this study was conducted with an upper middle class population of students. The question of generalizability remains present: Would results remain consistent in a district with fewer resources and larger numbers of minorities or lower achieving students? Second, the authors of this study excluded students in special education. It would be a crucial piece of information to find out how students in special education would fare when "included" as opposed to isolated into self-contained or resource room settings. Finally, with the use of the stanine scores, the authors were unable to measure the effects of de-tracking on students at the highest and lowest levels of initial achievement (top and bottom 5%). A question still unanswered is would this specific group of students benefit from de-tracked math curriculum? With continued research,

these questions could be answered, as school leaders and officials continue to need guidance in making data-based decisions.

In June of 2008, Burke and Sass conducted research analyzing the influence of classroom peers on individual student performance in Florida Public Schools. Because of current policy issues such as the influence of school choice programs, ability tracking within schools, and mainstreaming of students in special education, the purpose of this study was to examine the potential for peers to affect individual student achievement and continue to clarify how the structure of peers could have an affect on achievement. Burke and Sass claim that unlike past studies, their data set was unique because it allowed them to identify each member of a given student's classroom peer group in elementary, middle, and high school as well as the classroom teacher responsible for instruction (Burke & Sass, 2008). The authors were able to control for individual student fixed effects, therefore eliminating typical bias that exists in these types of studies.

Using the Florida Comprehensive Assessment Test (norm-referenced) and the Sunshine State Standards Assessments (SSS) (criterion-referenced test) for Grades 3-10, Burke and Sass (2008) conducted a quasi-experimental study by controlling for "extremities operating through fixed peer characteristics" (also as known as exogenous effects) as much as possible (p. 22). The data sample covered five years of schools, 1999/2000 to 2003/2004, and included all public school students in the state of Florida. The sample was divided into three groups: elementary school (Grades 4 and 5), middle school (Grades 6, 7, and 8), and high school (Grades 9 and 10).

On examination of the data, the authors of this study found positive and significant peer effects within every level of schooling for both reading and math. It is important to note that this

effect is generally small; for every one point increase in the mean peer fixed effect, the individual experiences an increase of 0.44 points in his or her current gain score. The coefficient is the smallest for elementary school reading ($b=.015$, $p<.05$) and the highest for middle school reading ($b=.069$, $p<.05$). With that said, "Elementary school results show that the lowest ranked students appear to receive the greatest benefits from having higher quality peers (.82 point boost to their math gain score for every 1 point increase), but middle ranked students also receive sizable benefits (.10 point increase under that same conditions)" (Burke & Sass, p. 16). The authors feel that these results provide a strong argument in favor of distributing top students relatively evenly across classrooms at the elementary level rather than isolating them from other students (2008). Plainly, if a school leader's goal is to maximize student achievement, the data show evenly mixed groups rather than ability-tracked groups. Additionally, mixed groups should be done early in elementary school, as the data shows the effects are greater in elementary school than middle and high school.

Finally, the authors continued this study by looking at the best model for distributing high, medium, and low skilled students within a classroom setting. There are three scenarios that were examined:

- Classroom 1 - 60% low, 30% middle, 20% high
- Classroom 2 - 10% low, 30% middle, 60% high
- Classroom 3 - 5% low, 90% middle, 5% high

The second experiment had the most desirable outcomes; the lowest students benefited by a large margin, the middle ability students benefited modestly, and the high students remained at status quo.

Results indicated that peer effects are only significant at the classroom level and not at the general grade level and that peer effects are not "one size fits all." The data show that the weakest students appear to experience the biggest positive influence from having higher quality peers, while high ability students appear to experience the weakest spillover from mean peer ability. When making placement decisions, the authors caution school leaders to place low ability students with their top quality peers, but in small increments, as too many low ability peers may fully offset the gains of the high performing peers and cause the opposite of desired effects.

Student Variables

Socioeconomic Status

Socioeconomic status (SES) is one of the most commonly used variables in education research. Researchers continue to use this variable to examine relationships between student achievement and their background/upbringing, hoping to draw conclusions and provide education leaders with answers and future policy implications. With that said, there is still much debate on the relationship between SES and academic achievement as new results are proving inconsistent, relationships ranging from a strong correlation to no significant correlation at all (Sirin, 2005). Additionally, recent empirical evidence from Sirin (2005) and Harwell and LeBeau (2010) has indicated that the generic term SES alone may not be an accurate way to examine and explain student achievement. Many researchers use SES and social class interchangeably, without any rationale or clarification, to refer to social and economic

characteristics of students. Instead, these researchers note that their studies have rendered a relationship between SES and student achievement, but "the relation is moderated by the unit, the source, the range of SES variable, and the type of SES achievement measure" (Sirin, p. 432). As stated by Duncan, Featherman, and Duncan (1972) "SES is defined by three main indicators which include the following: parental income, parental education, and parental occupation" (p. 87). Bollen, Glanville, and Stecklov (2001) and Hauser and Huang (1997) further clarify this definition by explaining that each indicator describes a substantially different aspect of SES that should be considered separate from the others. It is not enough to use a variable defined as "SES"; this variable must be defined by unit, as these new findings and limitations may alter the reliability of a study.

In 1966, James Samuel Coleman was commissioned by the U.S. Department of Education to conduct the landmark study entitled *Equality of Educational Opportunity,* most commonly referenced as the Coleman Report. This report was one of the largest in United States history, with more than 150,000 students in the sample and over 700 pages of report containing an array of information detailing: school environment (i.e., school facilities, services, curriculum, staff, and fellow students), pupil achievement and motivation (i.e., outcomes of schooling, integration and achievement), future teachers of minority groups, higher education, non-enrollment records, case studies of school integration, and special studies, among other various findings However, the most controversial was the discovery that once SES was controlled for, school resources have very little influence on academic performance (Gamoran & Long, 2006). Findings of the Coleman Report indicated that student background and SES were more important in determining educational outcomes than per pupil spending (Coleman,

Campbell, Hobson, McPartland, Mood, Weinfeld, & York, 1966). More specifically, results indicated the following:

> Socioeconomic status explained a greater proportion of student test scores than other measures of school resources such as class size and teacher characteristics; 49% student background, approximately 42% teacher quality, and 8% class size. The report showed that a school's average student characteristics, such as poverty and attitudes toward school, often had a greater influence on student achievement than teachers and schools, and that the average teacher characteristics at a school had a small influence on a school's mean achievement (p. 29).

The findings of Coleman caused studies to continue focusing on the influence of socio-economic status on student achievement. Fory years after the initial Coleman Report, Gamoran and Long (2006) attempted to synthesize the research conducted and draw some conclusions in the following areas: (1) examination of the main findings of EEO (also known as the Coleman Report) and determine whether they still hold and are accurate today, (2) reassessment of the results of the Coleman report on an international scale, (3) implications of the Coleman report for the debate of school vouchers and school choice, and (4) changes over the last 40 years in educational opportunity and equality and their influence on current education policy.

In examining racial segregation, Gamoran and Long (2006) explain that from 1954 through the mid 1970s, legal segregation was eliminated and Black-White school segregation dropped dramatically. Although this decline was expected to continue, the gains in desegregation peaked in the 1980s and then reversed in the 1990s by de facto segregation and the urban-suburban population shift (Gamoran & Long, 2006, p. 14). Additionally, the courts declared that school systems move from a "unitary status," meaning that schools are no longer

segregated by their own actions. This change caused desegregation programs to be dismantled and more schools to inadvertently become segregated. As a result, the proportion of Black students enrolled in predominantly minority schools has returned to the time of Coleman.

Another important area that Coleman examined was the Black-White achievement gap. Coleman et al. (1966) found that 85% of Black students who received an education through the twelfth grade scored below the average for Whites. Gamoran and Long (2005) emphasized "on average, Blacks scored a standard deviation below White students in academic achievement" (p. 5). Since the Coleman Report, the NAEP (National Assessment of Educational Progress) shows that the gap has narrowed from 1.2 standard deviations in 1971, when 17-year-old Black and White children's standardized test scores were examined, to 0.69 by 1996. Mathematics saw similar results from 1.33 to 0.89 standard deviation units. Unfortunately, this has not been a steady trend as the numbers have fluctuated but have ended favorably, a 27 point difference in reading scores in 2004 (Perie, Moran & Lutkus, 2005).

In examination of the evidence, Gamoran and Long (2006), supported by the results of Coleman, Kelly, & Moore (1975), found that although not conclusively documented there is a link between school desegregation and academic achievement. It was also found that peer composition held a modest significant influence; Black students who had more White classmates tended to score higher. Grissmer, Flanagan, and Williamson (1998) refuted this evidence, stating that desegregation has not been a prominent source of change in the achievement gap.

Finally, the most controversial finding of the Coleman report was that school resources had little effect on educational outcomes once family background was controlled for (2005)." Other researchers decided to conduct their own analysis to determine if their findings would

replicate those of Coleman (1966). Averch, Carroll, Donaldson, Kiesling, and Pincus (1974) discovered inconsistencies when attempting to identify which school resources dominated the influence on student achievement. Averch et al. reported mixed results; however, they did arrive at a similar conclusion to Coleman et al. that a student's socioeconomic background is the largest contributor to student success (Gamoran & Long, 2006, p.7). Gamoran and Long (2006) also highlighted studies that challenged the findings of the Coleman Report, which Gamoran and Long summarized:

> These critiques have included arguments that Coleman's cross-sectional study could not adequately capture causal effects, that Coleman assumed a linear and additive relation between resources and learning, that cross-sectional measures of reading achievement could not distinguish between learning that occurs at home and learning that occurs at school, and that Coleman's estimation of school effects by measures of percentage of variance explained were sensitive to assumptions about causal ordering (p.7).

Following this study, Jencks et al. (1972) agreed that there was value in the Coleman Report findings that determined little variance in resources from Black and White schools across the United States; however, Jencks et al. also found significance in other results, such as the academic achievement increase of students with lower socioeconomic background that attended schools with affluent peers. Jencks et al.'s investigation determined that after measures were taken into account for "sampling procedures, information-gathering techniques, and analytic methods," the Coleman Report results "[held] up surprisingly well" (p. 70). Additionally, Jencks et al. (1972) found that family background had a strong effect on student performance, noting that until inequalities pertaining to occupational status, education and parents' income are addressed, inequalities will continue to exist in educational institutions.

Controversy regarding the Coleman Report continued into the 1990s as researchers investigated the relationship between student achievement and school resources. Greenwald, Hedges, and Laine (1996) staged an ongoing battle with Hanushek regarding research done in this area. Greenwald et al. (1996) explained that "Hanushek's synthesis method, vote counting, consists of categorizing, by significance and direction, the relations between school resource inputs and student outcomes (including but not limited to achievement)" (p.362). Greenwald et al. (1996) explained that the methods employed by Hanushek were outdated and unreliable.

Greenwald et al. (1996) determined to prove his results correct, "that the data on the relations between school resource inputs and student outcomes, including achievement, were substantially more consistent and positive than he believed," conducted a meta-analysis using the following criteria for inclusion:

1. The data are presented in a refereed journal or a book.

2. The data originate in schools in the United States.

3. The outcome measure is some form of academic achievement.

4. The level of aggregation is at the level of school districts or smaller units.

5. The model controls for socioeconomic characteristics or is either longitudinal (including a pretest and a posttest) or quasi-longitudinal (including IQ or a measure of earlier achievement as an input).

6. The data are stochastically independent of other data included in the universe. (pp. 364-365)

Greenwald et al. (1996) concluded that their meta-analysis confirmed "that school resources [were] systematically related to student achievement and that these relations [were] large enough to be educationally important" (p. 394). Greenwald et al. emphasized that although

they anticipated their findings would validate a relationship between resources and student achievement, they were surprised that the "conclusions |were| so uniform in direction and comparable in magnitude" (p.385).

Sirin (2005) continued research in the area of the influence of SES on student achievement by conducting a meta-analytic review of research and journal articles between 1990 and 2000. This review was a replication of White's 1982 meta-analytic study, which focused on empirical studies published before 1980 examining the relation between SES and academic achievement. In examining the relationship between SES and achievement, White (1982) found that relation varies significantly with a number of factors such as the types of SES and academic achievement measures. Since White's 1982 analysis, a plethora of new empirical studies have explored the same ideas and the new results are inconsistent. The results range from a strong relation (Lamdin, 1996; Sutton & Soderstrom, 1999) to no significant correlation at all (Ripple & Luthar, 2000; Seyfried, 1998). Sirin attempted to review the more current works as he surmised that SES is one of the most widely used contextual variables in education research and as such must be documented and its limitations noted accordingly. Sirin aimed to (1) determine the magnitude of the relation between SES and academic achievement, (2) assess the extent to which this relation is influenced by various methodological characteristics (the type of SES or academic measure) and student characteristics (grade level, ethnicity, and school location), and (3) replicate White's previous work with recent data.

The sample for this meta-analysis consisted of 101,157 students from 6,871 schools, 128 school districts, and 74 independent samples in studies that were conducted from 1990-2000. Sirin used the following criteria to select research to be included in this study:

1. Apply a measure of SES and academic achievement

2. Report quantitative data in sufficient statistical detail for calculation
3. Include in its sample students from Grades K-12
4. Be published in a professional journal from 1990-2000
5. Include in its sample students from the United States

Sirin (2005) uses a sound methodology, as he is descriptive in his steps. Overall, 75 independent samples from 58 published journal articles were used. This translated into sample size groups of 26 to 21,263 with a mean of 1,580.58, a median of 367.5 and a standard deviation of 3,726.32. Sirin found that the majority of studies had multiple indicators and variables of interest, such as student characteristics and SES; 207 correlations were coded, ranging from .005 to .77.

The samples with the student-level data averaged ES (effect size) for the fixed effects model was .28 with a 95% confidence interval of .28 to .29, and it was significantly different from zero ($z = 91.75$, $p <.001$). The average ES for the random effects model was .27 with a 95% confidence interval of .23 to .30, and it was significantly different from zero ($z = 14.26, p < .001$). For the samples with the aggregated level data, however, the correlations ranged from .11 to . 85,with a mean of .60(SD= 22). The weighted ES ranged from .11 to 1.25. The average ES for the fixed effects model was .67 with a 95% confidence interval of .66 to .67, and it was significantly different from zero ($z = 147.56$, $p < .001$). The average ES for the random effects model was .64 with a 95% confidence interval of .57 to .70, and it was significantly different from zero ($z = 13.27, p < .001$) (Sirin, 2005). These results indicated that as the variables became more prominent in the study, the effect size increased, thus indicating a significant difference between the comparison groups.

Within the studies used for this meta-analysis, six SES components were used to assess SES, which included the following: parental education (k=30), parental occupation (k=15), parental income (k-14), and eligibility for free and reduced lunch programs (k=10). "A weighted ANOVA revealed that the average ES was .28 for parental occupation, .29 for parental income, and .30 for parental education. SES measures based on 'home resources' produced the highest mean ES (.51), followed by eligibility for free or reduced lunch programs (.33)" (Sirin, p. 426).

Sirin (2005) compares the findings of this study to the meta-analysis of Lipsey and Wilson (1993) in that "of all the factors examined in the literature, family SES at the student level is one of the strongest correlates of academic performance" (p. 445). Sirin also indicates that at the school level, the correlations were even stronger. In other words, the review's overall finding indicates that a parent's location in the SES structure has a strong influence on students' academic achievement because it has a direct control on a parent's ability to provide resources at home and indirectly, providing the social capital needed to succeed in school (Coleman, 1988; Sirin, 2005). Consequently, family SES determines the type of school, community, and classroom environment to which the student has access (Reynolds & Walberg, 1992).

More importantly, the research by Sirin takes Coleman's work a step further because his findings indicate that the magnitude of the relationship between SES and achievement is contingent upon several factors, including the following: student's grade, minority status, and school location (2005). In terms of grade level, which results differed from the work of White (1982) and Coleman (1966), the overall trend showed that the SES achievement relationship increased significantly by each school level starting at primary and continuing through middle

school (2005). This provides evidence that the gap between low and high SES students is most likely to remain the same, if not to widen as students continues through school.

It was also found that SES was a stronger predictor of academic achievement for White students than for minority students. These results rendered that the more minority students in a sample, the weaker the association between SES and achievement. In turn, Sirin (2005) suggests that the neighborhood and school may "exert a more powerful effect on academic achievement in minority communities, specifically African American communities" (p. 436). This stems from work conducted by Dornbusch, Ritter, and Steinberg (1991), who discuss the effects of a minority community on academic achievement and state that "with African Americans in particular it is not solely because of their minority status but partly because fewer Whites live in neighborhoods with higher educational risk factors" (p. 29). Additionally, Sirin explains that these ideas also coincide with school location. Results of this study have shown that the influence of family SES varies for individuals depending on where they live and the cohort with whom they go to school (Sirin, 2005).

Although Sirin (2005) conducted a meta-analysis with a large diverse sample and disclosed proper research techniques, it is important to note the limitations to his work. First, this study was limited to works between 1999-2000 and included only studies with students enrolled in schools in the United States; therefore, it is possible that some studies were missed or not included on the basis that meta-analysis studies depend on the quality of work published by others. Also, the author notes that caution should be taken with the results because according to Lipsey and Wilson (1993), there are limitations in effect sizes of published versus unpublished studies, and it is possible that the results of this study are overestimated.

As mentioned by Sirin (2005) in his section entitled "Implications for Future Research," more studies need to be conducted to investigate specific categories of SES. Harwell and LeBeau (2010) investigated the most commonly used form of SES in educational research, student eligibility for free and reduced lunch (FRL). Because this study will use the FRL variable as an input variable, this information was included. Harwell and LeBeau attempt to clarify the FRL variable, explain its biases and deficiencies, and explain why, although it is continuously criticized for being a weak measure of SES, it continues to be used in research among educators.

The FRL variable is defined by criteria created by the federal government. The first criterion relies on income information provided by the householder. According to Harwell and LeBeau (2010), "A student is eligible for reduced lunch if their household income is 185% of the federal poverty guidelines and eligible for free lunch if their household income in less that 130% of the poverty guidelines" (p. 127). In other words, in 2008 a family had to have a household income less than $39,220 for reduced lunch and less that $27,560 for free lunch.

Another way that families are discovered as eligible for FRL is based on whether a household receives food stamps, has foster children, or participates in a federally funded assistance program. In this case, no applications need to be filed, as social service agencies directly involved with the families identify children who quality for this program.

The authors disclose that demographics of the population of students who receive FRL are scarce; however, using an ERS report from 2006, which included data sets from The National Health and Nutrition Examination Survey and the Public Elementary/Secondary School Universe Survey, the participation rates were found highest for elementary school children and declined

each year after (Harwell & LeBeau, 2010). Also, the participation of White, African American, and Hispanic students in the FRL program was proportionally equal, although White students were most commonly in the "reduced" category while African Americans and Hispanics were most likely in the "free" category. Results of the study indicate that schools with more African Americans have more students eligible for free and reduced lunch and that schools in urban areas tend to have more students eligible for FRL. Schools in the suburbs or rural areas tend to have fewer students eligible.

Harwell and Lebeau (2010) found that of all the SES measures available, free and reduced lunch eligibility was the most likely to be used because it provides easy access, is inexpensive, and requires minimal responses from participants in the sample; however, it is not valid as an indicator of access to household resources. With that said, the authors recommend that an important practice for education researchers is to "adopt and carefully describe what SES is intended to represent in his/her study and to provide a clear rationale for selecting that measure of SES" (p. 127).

In summary, Coleman reported in 1966 that the greatest influence on student academic achievement was socioeconomic status. As meta-analyses continue to be conducted, Coleman's initial research continues to be supported (Dornbusch, Ritter, and Steinberg, 1991; Gamoran & Long, 2006; Sirin, 2005). In fact, these findings are now beginning to be looked at specifically by category (Harwell & Lebeau, 2010; Sirin, 2005), as some aspects of SES have a stronger influence than others. As this research continues to develop and debate regarding which aspects of SES have the strongest influence on student achievement, one aspect remains consistent: SES is the single strongest predictor of student achievement. After reviewing the literature in this

area, the combination of SES and placement in an inclusion classroom may cause an even greater influence on the student achievement of general education students; therefore, this variable is critical to investigate the influences on various demographics of students within the general education population.

Student Attendance

Since the Coleman study was released in 1966 (Coleman et al.), social science researchers, specifically in the area of education, have continued to examine the demographic variables or "input factors" that have an effect on student achievement and school performance. Besides SES, which empirical evidence has continually proven to be the strongest predictor of student achievement, researchers Caldas (1993), Roby (2004), Lamdin (2001), and Chen and Stevenson (1995) have conducted empirical studies examining the relationship between student attendance and student achievement. Each of these empirical studies provides statistical evidence that demonstrates that a lack of attendance in school negatively influences student achievement as measured by standardized assessments. For the purposes of this study, attendance as a variable must be explored because it may provide evidence showing whether regular education students with poor attendance placed in an inclusion classroom may demonstrate different levels of achievement.

Caldas (1993) conducted a study that examined the direct effects on and contribution of several input factors on public school achievement in Louisiana. Caldas believed that although there are a lot of input variables that cannot be controlled, there are some structural factors that, if found to have a significant influence on student achievement, could be manipulated. One of these variables that could be manipulated was student attendance. Change/manipulation of this

input variable could occur through suspension/expulsion policies, incentive programs, attractive course offerings, and attendance policies (1993). Caldas designed a study using the data made available by the Louisiana State Department of Education looking at how much relative influence do input and process factors have over school achievement.

The sample for this study was 1,301 public schools classified as either an elementary, middle, or high school. Caldas notes that all alternative, vocational, and special education schools were excluded. Using a multiple regression analysis, Caldas defines the output variable as a combination of scores on Louisiana criterion-referenced and norm-referenced tests scores. In order to increase validity of the data, he combined scores on the annual state tests into one category called ACHIEVE. Although Caldas examines a variety of input variables (a combination of SES variables and demographic variables), for the purposes of this literature review only the results of the effect of student attendance on academic achievement is noted. Student attendance (PCTATTEN) was defined as the average daily student attendance divided by average daily student membership multiplied by 100.

Results of this study showed that the input factor with the strongest effect on ACHIEVE was percentage of student attendance ($ß=.187$, $p<.001$), although the SES input variable was 2.5 times greater than PCRATTEN. Additionally, in examination of a second model looking at elementary versus secondary schools, the magnitude of the effect of student attendance on achievement, although positive at both grade levels, was twice as strong in secondary ($t=.270$, $p<.001$) as in elementary ($t=.107$; $p<.001$). Caldas also states that it was specifically the percentage of student attendance that accounted for almost all of the additional variance of secondary school achievement explained by process factors; percentage of attendance was the

most important process variable in every model. As a result of this study's findings, Caldas suggests that efforts to increase secondary school attendance is likely to be rewarded in terms of student achievement.

Although Caldas (1993) provides some initial insight on the influence of student attendance on academic achievement, he neglects to calculate effect sizes for each of his variables, which raises some questions with his results. Additionally, Caldas does not make note of his limitations, which include a minimal explanation of his population as well as how many of sample members were excluded or had incomplete data.

Lamdin (1996) continued the work on student attendance by commenting that the past research that has been done has not used student attendance as an independent variable; therefore, using data from the Baltimore public elementary schools, he conducted a study using the production-function approach in which student attendance was considered an independent variable.

The sample for this study was collected by using data from 107 public elementary schools, using only schools with Grades K-5. Ten schools were excluded because they were not comparable to the K-5 school. Results were calculated using the dependent variable of the California Achievement Test (CAT) in both reading and mathematics. Additionally, Lamdin (1996) used the independent variables of SES or the percentage of students who do not qualify for free and/or reduced lunch, race, pupil teacher ratio, and student attendance. Although Lamdin did render results on a variety of variables, for the purposes of this study only data regarding student attendance will be reported. Using a regression analysis, Lamdin found the coefficient on the attendance variable was both positive and significant ($p<.05$) for a one-tailed

test in all nine of the specifications and significant at ($p<.01$) for eight of the nine. It is also important to note that in his statistical analysis, Lamdin fails to calculate an effect size for his variables. The results strongly suggest that "attendance does have a positive influence on student performance when other factors are held constant" (p. 158).

In his concluding statement, Lamdin (1996) does warn consumers to use the results with caution because it is difficult to isolate and control all the factors that could effect input variables such as school resources, the home life of each student, or other demographics that could influence the results. In fact, Borland and Howsen (1998) published a rebuttal to Lamdin's work suggesting that "Lamdin's results are biased because of his failure to include measures of student innate ability and competition in the explanation of student performance" (p. 196). To support this claim, Borland and Howsen conducted a study similar to Lamdin's, using a similar model; however, they used a different data set and included in their model a measure of student attendance, the pupil-teacher ratio, and expenditure per pupil. Results of their study found that the signs for student attendance were positive; however, they were not found statistically significant. In other words, Borland and Howsen (1998) concluded that student attendance will not lead to an increase in student performance.

In contrast to the work conducted by Borland and Howsen (1998), Roby (2004) conducted a study looking at the relationship between student achievement and school attendance as measured by the Ohio Proficiency Test. His interest stemmed from the work of Lamdin (1996), King (2000), and Johnston (2000), who concluded that the positive influence of good school attendance on academic achievement may be greater than historically thought." Additionally, Dekalb (1999) noted that student achievement continues to be affected in a

negative way by absenteeism which is supported by the work of Robins and Ratcliff (1978) who did a study using African-American males. Results of that study concluded that of the students consistently truant from elementary and high school, 75% did not graduate (1978).

Chen and Stevenson (1995) also performed a study that examined the mathematics achievement of Asian-American students from a cross-cultural perspective. The researchers found that attendance was an "achievement-related behavior" that influenced student achievement outcomes on exams; the more often a student was absent, the more poorly he or she performed on the exam With these conclusions, Roby (2004) attempted to add to the current body of research by examining if there was a positive and significant relationship between student attendance (school building averages) and student achievement as measured by the Ohio Proficiency Test.

Roby (2004) used the Pearson's r correlation statistic to analyze the relationship between student achievement and attendance. Although Roby examined Grades 4, 6, 9, and 12, only Grade 4 will be reported as this study focuses exclusively on students in Grade 4. Roby notes his used of the standard measures of central tendency as well as an independent t test to determine statistical significance. Data used for this study were taken from the Ohio Department of Education web site and information was used up to 1999. The sample was made up of 3,171 students, and 1,946 students were in fourth grade.

Results of this study show that there is a moderate positive relationship between student attendance and student achievement in Grade 4 ($r=.57$, $r2=.32$, $p<.01$). Additionally, the correlations were considered statistically significant at the $p<.01$ level. When the coefficient of determination was calculated for fourth graders ($r2=.32$), it shows that 32% of the variance held

in common with student achievement results relates to student attendance. When the top 10% and the bottom 10% of fourth grade students in Ohio public schools were examined, Roby (2004) found that the fourth grade comparison indicates a large variance ($t=9.70$) and was found statistically significant at the $p<.05$ level ($t=9.70$).

The statistics of this study indicate that not only do students who have higher attendance rates score higher on the Ohio Achievement Test, but the buildings with the higher student attendance rates have higher test scores averages. In other words, this study demonstrates that there is a statistically significant relationship between student attendance and student achievement in Ohio at the fourth, sixth, ninth, and twelfth grade levels, and the correlation is positive and moderate to strong. From the findings of this study, Roby (2004) finds the need for incentive programs to be implemented and for students to attend school consistently on a daily basis, or it will result in a significant loss of instructional time, which will decrease student achievement measured by standardized assessments.

Roby (2004) notes that his study does have limitations and should only be used as a resource for discussion and debate. Roby mentions that there are other variables that could play an important role in the correlation such as SES, aptitude, family make-up, student age, relationships with teachers and value in school. Although these variables were not examined individually, they also may influence the given results. Roby (2004) also adds a rebuttal to Borland and Howsen (1998) and explains, with the use of a new data set, that the results of this study support those of Lamdin (1996) in that there is a relationship between student achievement and student attendance.

Examination of the empirical evidence clearly establishes a statistically significant relationship between student achievement and attendance. Empirical evidence exists confirming that as a student's absenteeism rate increases, their academic performance worsens. Thus, attendance is a strong predictor of student achievement on state-mandated assessments and should be included as a variable for continued research.

Student Gender

Another variable that is commonly explored when analyzing variables influencing student achievement is gender. Although a variety of studies continue to use gender as an influence variable, recent research has found that there are underlying factors which cause differences in achievement between gender groups. Wilkins, Zembylas, and Travers (2002) documented that personal, instructional, and environmental factors account for gender discrepancies. According to research, these individual factors include socioeconomic status (Drukker et al., 2009), culture and surroundings (Pajares, 2002), neurological composition (Gurian & Stevens, 2004), policy and regulations contrived by state and federal agencies (Gunzelmann & Connell, 2006), as well as biology (Salamone, 2003). With that said, there is no empirical evidence supporting the gender gap or a continued growth in the academic performance of males and females.

In 1997, Warren W. Willingham, scientist for Educational Testing Service, and President Nancy S. Cole conducted a four-year study analyzing gender differences on assessments. Willingham and Cole (1997) exposed several myths surrounding the conventional notion that girls generally do well in the liberal arts, whereas boys tend to excel in mathematics and the sciences. Willingham and Cole concluded that data revealed that there was essentially no

difference between females and males for 74 assessments at the twelfth grade level across 15 subject areas. The gender gaps of the 1960s have since narrowed. The researchers refuted the belief that boys outperformed girls in mathematics and science. The authors asserted that gender differences do not necessarily account for gaps in student achievement. In fact, it is the individual factors and personality traits that have a stronger influence on the academic achievement of students.

In 2008, Marks analyzed the 2000 Programme for International Student Assessment Project (OECD, 2001) results to determine how achievement in reading and math was influenced by student gender. The sample for this study included student achievement scores for over 172,000 15-year-old students in 6000 schools in 32 countries.

Results of this study found that "the gender gaps in reading and mathematics are highly correlated and that the magnitude of the gaps reflects the implementation and success or otherwise of policies designed to improve girls' educational outcomes" (Marks, 2008. p. 106). It is also important to note that in 15 of the 31 countries the gender differences in mathematics were not found to be statistically significant. Additionally, in many OCED countries until recently there were sizable gender gaps in mathematics with girls performing less well than boys. Similarly, the lower the educational and occupational expectations of girls, the more the gender inequalities reflected academic outcomes (Marks, 2008). However, this study shows that fundamental changes have been made; the gender gap is continuing to decline as policies reflect equal education for all.

As noted by Marks (2008) many policies in a variety of countries throughout the world are finally placing equal emphasis on the education of both males and females; however, it is not to be overlooked that although policy continues to modernize, there is still an under-

representation of women in the science, technology, and mathematics fields. Although evidence exists to contradict this point, this rumor continues to circulate, indicating that males outperform their females counterparts in math and science achievement and that stereotypes exist about female inferiority in mathematics (Else-Quest, Hyde, & Linn, 2010; Hedges & Nowel, 1995; Hyde, Fennema, & Lamon, 1990; Hyde, Lindberg, Linn, Ellis & Williams, 2008).

In order to examine this common misconception, Else-Quest et al. (2010) conducted a study looking at gender differences cross-nationally. Analyzing the 2003 Trends in International Mathematics and Science Study and PISA results, the researchers determined that "on average, male and females differ very little in mathematics achievement, despite more positive math attitudes and effects among males" (p. 125).

Using a meta-analysis to estimate the magnitude of gender differences in mathematics achievement, the data set for the TIMSS 2003 included 46 countries and a total of 219,612 students. The TIMSS included the following content area domains: numbers, algebra, measurement, geometry, and data. Additionally, the PISA 2003 was used because of its focus being predominantly mathematics. The PISA sample included 41 countries and a total of 273,883 students. To analyze the robust sample, the authors "conducted a meta-analysis on each composite and content domain of the TIMSS and PISA data sets as well as computing effect sizes of gender differences for each of the 11 measures of mathematics achievement (Else-Quest et al., 2010).

Results of the TIMSS 2003 indicated that boys and girls performed similarly overall in the content area domains of Numbers, Algebra, Data, Geometry, and Measurement (k=46, d=-.01). In other words, the effect size reflects a gender difference of less than 1 point on the

TIMSS, which had an international score of 467. On the PISA-Math, results showed that boys performed slightly better than girls overall; however, the effect size of d=.11 explains that the gender differences are very small in magnitude. Else-Quest et al. (2010) found that the results of this study in comparison to previous data show that the average effect size has changed very little. As in the work of Baker and Jones (1993), (d=0.03) and Mullis, Martin, Fierros, Goldberg, and Stemler (2000), (d=0.08), the effect sizes were small and there was a consistent pattern of gender similarity over two decades of research. It is also important to note that for over half of the nations the gender gap has remained near zero or has even decreased in magnitude to become negligible.

On examination of the evidence, it is clear that although there was once a large achievement gap between males and females, that gap has become small to almost non-existent. In fact, research in this area is still mixed and has shown that differences in student achievement are influenced more by individual student factors such as attitude, SES, and family. For the purposes of this study, gender will be included as a variable given the variety of empirical evidence supporting that student gender has a statistically significant effect on student achievement.

Theoretical Framework

As research continues to be conducted in the area of special education, a multitude of variables are examined in order to clarify which has the most significant influence on student achievement. It can be said that these variables, including socioeconomic status as measured by free and reduced lunch data, student attendance, peer influence, placement in an inclusive classroom model, and teacher experience, all have a research base which provides evidence that

they have an influence on student achievement. With that said, each of these variables can be grouped into three distinct "input" categories, including the following: student, school, and teacher variables. For the purposes of this study, each of these categories has been shown to have an influence on the "output variable" of student achievement on the NJ ASK 4 in language arts and mathematics.

Production function theory, which serves as the theoretical framework for this study, is defined as describing the relationship between school inputs and student outcomes (Greenwald, Hedges, & Laine, 1996). As stated by Caldas (1993), "The factors that affect student achievement over which school officials have little or no control have been termed *input factors*, whereas those factors over which school officials do have control have been termed *change or process factors*" *(p. 224).* In other words, gender, socio-economic status, and student attendance would be considered input factors because they are comprised of demographic and socioeconomic characteristics that can have an influence on student achievement but cannot be altered in their entirety or easily by school officials. Other variables known as *process variables*, which can be changed or altered, include the following: placement in an inclusive setting, peer influence, teacher experience, and degree.

Although research on input variables such as SES, gender, and attendance continue to be conducted, researchers continue to yield mixed results. For example, Coleman (1966) found that school resources had a surprisingly small influence on student achievement, while student SES accounted for the majority of the variance in student achievement. Coleman's results are both supported and yet still debated, especially by Hanushek (1979), who claims that Coleman assumed a linear and additive relation between resources and learning and that Coleman's

estimation of school effects by measures of percentage of variance explained was sensitive to assumptions about causal ordering (p. 7).

Regardless of the continuing debate, it can be concluded that these variables, both input and process, individually or collectively, have an influence on student standardized test scores or potentially support a plausible change in academic behavior. The question that still remains is how does the variable of placement in an inclusive setting influence the student achievement of general education students? Furthermore, when "inclusion" is examined from a theoretical lens, there is not one concrete theory or foundation that can be uncovered. Most of what is revealed stems from court-based decisions, past incidences such as Willowbrook State School, and legislation such as NCLB and IDEA promoting equality. However, after extensive research, a variety of theories were discovered that, when combined, attempt to explain not only a theoretical understanding of inclusion but why inclusion is essential for all children and may or may not affect achievement of students placed within that setting. Plato and Aristotle would both argue that "in order to be who we truly are, we must live in the 'true' society comprised of natural proportions (Kraut, 2012, p. 13).

Ethical Components of Inclusion

Even earlier than the nineteenth century, philosophers and theorists had examined a variety of issues regarding equality, social justice, and human nature. Although the debate regarding *inclusion* has only sparked nationwide attention over the last 20 years and is seen as predominantly a legal issue through NCLB and IDEIA, the building blocks were formed long ago, calling for the inclusion of all citizens and the realization that isolation was not the answer to diversity and disabilities present within a society.

Throughout history, people perceived as being "different" were vulnerable to harsh practices such as isolation, slavery, physical abuse, and societal abandonment. At one point it was even believed that disabilities were caused by hereditary factors that, if left unchecked, would result in widespread social problems (Gelb, 1995). As a result, many doctors attempted to "cure" this so called "deviance" and laws were passed promoting sterilization of "deviant individuals," prohibiting them from marriage, having children, and living in homes with "normal individuals." The government created "institutions" (also known as schools, hospitals, colonies, prisons, and asylums) to house these individuals in isolation away from society.

Beginning in the nineteenth century and continuing throughout the twentieth century, these institutions continued to grow. In an effort to maintain order and control, personnel employed in these establishments implemented strict rules and regulations forcing the individuals to wear institutional clothing, identification tags, identification numbers, live in locked units with bare walls, and limiting outdoor recreation activities. As a result of this inhumane treatment, funding for these programs decreased as patients were seen as uncured and classified as "terminal." Sadly, this situation was unchanged for nearly five decades and continued on a downward spiral through the Depression in the 1930s. By the 1950s, more than 500,000 persons were committed to mental hospitals throughout the United States, and comparable numbers of persons with mental retardation lived in segregated institutions (Hardman, et al., 2002).

For the better part of the twentieth century, families who had a child with disabilities were unable to get help for basic needs such as medical and dental care, social services, and education. In response to the lack of government support, parents begin to organize into

organizations such as the National Association for Retarded Children in 1949 to advocate for the rights of persons with disabilities. Alongside the Civil Rights Movement, which occurred during the 1950s and 1960s, the spark that began the fight for the equality of African Americans concurrently sparked the "Inclusive Movement," in which advocates protested equal rights for individuals with disabilities. According to the Constitution of the United States, civil rights are the basic legal rights a person must possess to secure a status of equal citizenship which encompass personal, political, and economic rights. Results of the Civil Rights movement found that these human rights cannot be denied to a person on the basis of race, color, sex, religion, or disability.

According to the early works of Plato and Aristotle, it is inconceivable that humans would deprive other humans of their right to be members of society, yet for centuries persons who were labeled as "different" were deprived of certain freedoms awarded by the Constitution of the United States. The concept of human nature explored by both Plato and Aristotle emphasizes that individuals must be exposed to a "true society" because who we are depends on what kind of soul we have and our soul develops through discovery. Plato believed that normative implications resided in success or failure and that success or failure in life depends upon what sort of society we live in. Human life needs to be political, spent in the discovery of the proper manner in which sociality ought to be organized and then in the practical implementation of that ideal in our own societies.

Examining the work of Plato provides valuable justification for inclusive practices. On one hand, inclusion means learning from others both academically and socially. If all disabled children learn only with other disabled people, it will result in an inability to interact with non-

disabled society, learning accepted norms and social standing. On the other hand, if all non-disabled students are educated in idealistic classrooms, they will not learn how to be role models, value "difference," or how to live in a "true society," a country founded on social justice and civil rights.

Not only were individuals with disabilities being deprived their natural rights, but they were being deprived basic legal rights as required by the U. S. Constitution. In 1971, John Rawls, an American philosopher, examined what he called the "theory of justice," which views justice as "fairness from an impartial point of view." These principles were the foundation for the legislation known first as The Americans with Disabilities Act (ADA) and eventually The Individuals with Disabilities Act (IDEA), which awarded legal rights to individuals with disabilities.

Rawls' theory of justice is based on certain guiding principles. The first principle guarantees the basic rights and liberties needed to secure the fundamental interests of free and equal citizens to pursue a wide range of conceptions of the good. The second principle provides fair equality of educational and employment opportunities, enabling all to fairly compete for the powers and prerogatives of office; and it secures for all a guaranteed minimum of the means (including income and wealth) that individuals need to pursue their interests and to maintain their self-respect as free and equal persons (Freeman, 1971, p. 21). Before the Civil Rights and Inclusive Movements of the 1950s, individuals with disabilities were denied their equal basic rights--including education, safety, and citizenship. Today, in order to "be just" and follow current laws that provide fair opportunities for all individuals both with and without disabilities,

school officials must create inclusive environments for disabled students that foster structural, academic, and social opportunities equal to their non-disabled peers.

Given that Rawls provides theoretical justification for an inclusive environment through justice, these environments must maintain proportional equality, as each disabled child deserves the right to be exposed (to the maximum extent appropriate) to his or her non-disabled peers. Antoin E. Murphy, a theorist who lived from 1601 to 1758, set the foundation for the need of proportional equality. Murphy examined proportional equality from a monetary point of view, supply and demand. Murphy realized that production had to be comparable to the demand that was required by its citizens. He concluded that proportional equality indicates that equal output is demanded with equal input. (Velde, 1999, p. 202). Proportional equality relates not only to a simple mathematical formula as used by Murphy, but to looking at a whole individual and deciding what is proportionally adequate for him or her in society. Structurally, inclusion promotes proportional equality insofar as it requires a proportionally natural environment in the classroom for both students with and without disabilities. Students are assigned to classes with consideration to the natural proportions of the population that live in the school's jurisdiction. For example, if 12% of the students have disabilities, then the school team would plan to place students in classes in as close to that proportion as possible. Generally, schools try to assign no more than 20% of the students with special education academic needs to a class at one time.

The proportion of all people with disability labels in the general population is about 13% to 15%. People with the most severe disabilities represent less than 1% of the general population. When students with disability labels attend their home school, there is generally a natural proportion represented. School buildings should consider the natural proportion when assigning

students to classrooms. Classrooms which consider the natural proportion will not have more than 15% of its members who have disability labels and no more than one of these students will have a label of severe disabilities. When the overall structure of inclusion is examined, it is clear that it exposes children to proportional equality and a "true society." Therefore, although inclusion can be seen simply as "law," there is a theoretical foundation for the creation of an environment/structure which allows equal access to society, resources, and opportunity for all individuals, free of discrimination and inequality.

Achievement Components of Inclusion

Although Plato, Rawls, and Murphy provide the theoretical backing from a structural and legal perspective, this study looks at inclusion in schools, specifically the influence of student placement in an inclusive environment on academic achievement. The work of Itard in 1799, followed by the theories developed by Sternberg, Piaget, and Bandura, provide theoretical insight as to why students with and without disabilities placed in an inclusive setting may render different output (academic achievement) results on state standardized assessments.

Early implications of the influence of inclusion on the education of individuals with disabilities began with a physician named Jean Marc Itard in 1799. Itard's work is reflected today in modern medical, psychological, social, and education intervention models, as he believed that environment in conjunction with physiological stimulation could contribute to the learning potential of any human being. The work of Itard was a further investigation of the earlier work done by Philippe Pinel (1742-1826), who advocated that people labeled as "insane or idiots" needed to be treated humanely, and John Locke (1632-1704), who described the mind as a blank slate that was open to all types of new stimuli (Hardman, et al., 2002, p. 17). Both

Pinel and Locke created the common debate of "**nature versus nurture,**" more specifically what role genetics and environment play in determining an ability to learn.

Itard tested the theories of Locke and Pinel in his study of Victor, the "wild boy of Aveyron." Victor was 12 years old, lacking language and exhibiting uncontrollable behavior, when found in the woods by hunters. Ignoring the diagnosis that Victor was an "incurable idiot," Itard took Victor into his home and put him in a program of sensory stimulation, language, and human interaction. Within five years of working with Itard and his housekeeper, Victor developed some verbal language and became more socialized. The result of Itard's study of Victor was landmark documentation that learning is possible even for individuals described as "helpless or incurable" (Hardman, et al., 2002, p. 17). This notion of how to educate and to teach was something that although it did not produce the effects hoped for (making Victor completely verbal), did prove to be a step towards new systems of pedagogy. Itard's ideas and theories would continue to be developed and investigated by future theorists.

Centuries later, Robert Sternberg's theory of "triarchic intelligence" viewed intelligence or learning as shaped from one's life and experiences. Psychologist Robert Sternberg (1985) defined intelligence as "mental activity directed toward purposive adaptation to, and selection and shaping of, real-world environments relevant to one's life" (Sternberg et al., p. 45). Sternberg proposed what he refers to as "successful intelligence," which is comprised of three different factors:

Analytical intelligence: This component refers to problem-solving abilities.

Creative intelligence: This aspect of intelligence involves the ability to deal with new situations using past experiences and current skills.

Practical intelligence: This element refers to the ability to adapt to a changing environment.

While there has been considerable debate over the exact nature of intelligence, no definitive conceptualization has emerged. Today, psychologists often account for the many different theoretical viewpoints when discussing intelligence and acknowledge that this debate is ongoing. As a result, placement in an inclusion classroom could be seen as an enhancement of "intelligence." Students both with and without disabilities are exposed to a variety of practical situations, which create new schema and the ability to adapt to not only a variety of people, but a variety of situations that would represent "real life." Baker (1994) in his meta-analysis found positive effect sizes from 0.08 to .44, meaning that students with disabilities educated in the regular classroom do better academically and socially than comparable students in non-inclusive settings (p. 34). From the perspective of Sternberg, this would result in an increase in intelligence, which could have a direct influence on student achievement on standardized tests.

Similarly, Piaget also viewed learning through social interactions and the building of new experiences in his combination of social constructivist theory and theory of cognitive development (Sternberg, 1985, p. 44). Social constructivist theory views each learner as a unique individual with unique needs and backgrounds. In other words, young children develop their thinking abilities by interacting with other children, adults, and the physical world. From the social constructivist viewpoint, it is thus important to take into account the background and culture of the learner throughout the learning process, as this background also helps to shape the knowledge and truth that the learner creates, discovers, and attains in the learning process. Furthermore, constructivist theory supports the learner as an active participant in constructing his

or her own knowledge and emphasizes the importance of the learner being an essential component of his or her own individual learning process. Piaget concludes that only when a learner is involved in learning does he or she continue to grow both socially and academically (Sternberg, 1985, p. 45).

Finally, Bandura's social learning theory plays a large part in educating students with special needs in inclusive education settings. Social learning theory was first looked at by Julian Rotter in the realm of social learning and clinical psychology in 1954 as he attempted to move psychology away from behavioral theory. Rotter initially suggested that "behavior is influenced by environmental factors or stimuli and not by psychological factors alone, as first thought by Skinner and Thorndike (Ormrod, 2008, p. 146). Bandura expanded on the ideas of Rotter as well as the work of Miller and Dollard. Social learning theory incorporates aspects of behavioral learning (assumes that environment causes people to behave in certain ways) and cognitive learning (psychological factors are important for influencing how one behaves). This theory is a combination of environmental and psychological factors that establishes three requirements on how people learn and model behavior: retention (remembering what one observed), reproduction (ability to reproduce the behavior) and motivation (reason to want to adopt the behavior).

One of the first and most important implications of social learning theory is that *students often learn a great deal simply by observing others* (Ormrod, 2008, p. 145). The implication is that students learn from watching their age-appropriate peers. Students with disabilities need to see how their peers interact with others, how they conduct themselves in various school environments, and how they respond to various unexpected situations. Without peers to exhibit and model appropriate behaviors, special education children would be more likely to duplicate

inappropriate behaviors that are displayed by other students in special education. Observing appropriate behaviors by means of age-appropriate models is an integral piece of justifying inclusion as well as developing appropriate learning behaviors.

Modeling is another educational implication for social learning theory. The regular education students enrolled in the inclusion environment become the role models, providing an alternative to shaping new teaching new behaviors (Ormrod, 2008). Students can serve as informal models, showing students proper behaviors in various settings, or they can be used as formal models or examples. A teacher can always show a special-needs student how a fellow peer is acting during a reading lesson and use that behavior as a model. This idea is one of the four "essential conditions that must exist in order to promote effective learning" (Ormrod, 2008, pg. 145). The other conditions include: attention, retention, motor reproduction, and motivation (Ormrod, p. 145).

In order for the essential conditions to exist, it is the responsibility of the teacher to instill in children *a belief of self-confidence,* which has numerous implications for the special education population. Ormrod (2008) discusses that students' self-esteem affects their learning and academic achievement In other words, a teacher must always promote a learning environment that fosters learning. Without the encouragement and support of the teacher, students will be less likely to develop the self-esteem and confidence they need to develop new ideas and achieve goals, which eventually turns into "generalized self-efficacy" (Ormrod, p. 147).

Conclusion

Figure 1 presents the final conceptual framework and the input/output theoretical framework used to guide the study. Inclusion is a practice that must take place in schools for

legal, ethical, and learning purposes in order to provide each child equal opportunity to a "real society" (Rawls, 1971). Researchers and school professionals continue to debate what inputs have a direct influence on student achievement, as the empirical evidence provides inconsistencies and mixed results as to the influence of placement in an inclusive classroom on the academic achievement of general education students. Much of the evidence does provide sound data supporting academic achievement for students with disabilities in inclusive settings; however, because of the lack of research available on general education students, this study attempts to add to the slim body of literature in the hope of clarifying this issue for policymakers as well as for school administrators.

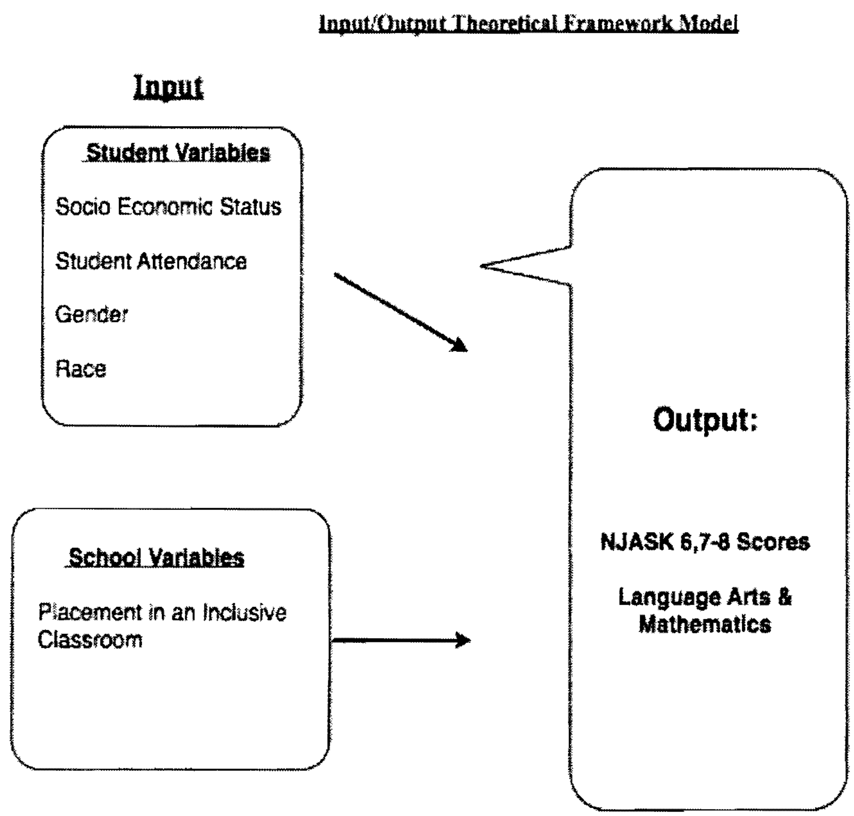


Figure 1. Input/Output Theoretical Framework.

Chapter III

METHODOLOGY

Introduction

The purpose of this quantitative study was to examine the influence of placement in an inclusion classroom on the academic achievement of general education students in Grades 6,7, and 8 as measured by the NJ ASK in both language arts and mathematics. This study aimed to produce research-based evidence regarding the education policy of "inclusion" and its academic influence on the general education student population in middle school. Much emphasis has been placed on the positive social and academic influences of students classified with disabilities; however, little is known about how general education students are influenced academically. With the use of a combination of student/school variables deemed to have an influence on academic achievement as well as the variable of placement in an "inclusion" setting, this study aimed to add empirical evidence to a limited and mixed body of existing literature, helping school administrators and policy makers to create policies and make decisions based on research based evidence.

Research Design

This study was conducted using an ex-post facto, correlational non-experimental design due to the lack of random assignment of subjects to the treatment and control conditions with quantitative methods. In order to determine which student and school variables had a statistically significant relationship to student achievement, I used simultaneous multiple regression models. This predictive model is used when the researcher is trying to determine how much of the variance in the outcome (dependent variable) can be explained by a group of

predictor (independent) variables. This method is typically used to explore and maximize prediction (Pedhazur, 1997). To support the work of Pedhazur, Rubinfeld (1986) explains that, "multiple regression also may be useful (1) in determining whether a particular effect is present, (2) in measuring the magnitude of a particular effect, and (3) in forecasting what a particular effect would be, but for an intervening event" (pp. 181-182). Additionally, examination of the literature identified the specific variables that affect student achievement; however, the extent to which these variables influence student achievement on the NJ ASK 6-8 is unknown. Because of this lack of knowledge, simultaneous regression was used. Researchers use simultaneous regression when they have a limited number of predictors and are unsure of which variables would create the best prediction equation model. Through the use of this methodology, the researcher was able to test stated hypotheses relating to the independent variables of placement in an inclusive/non-inclusive setting, student SES, attendance, and race on the dependent variable of non-disabled students' academic achievement.

Sample Population

The participants from this study were selected from an urban lower middle class PreK-12 school district located in central New Jersey. According to the Census Bureau, this city has a population of approximately 39,394 people, 15,052 households, and 10,084 families. The racial makeup of the city is approximately 66.08% White, 22.80% African American, 0.14% Native American, 2.35% Asian, and 4.88% from other races. Hispanic and Latino are about 14.40% of the population. The median income for a household in the city was $46,345, and the median income for a family was $54,903. Males had a median income of $39,457 versus $30,395 for females. The per capita income for the city was $21,314. About 5.0% of families and 6.4% of the

population were below the poverty line, including 8.1% of those under age 18 and 7.8% of those age 65 or over.

This Title I city district houses eight Pre-K-5 elementary schools, two Grade 6-8 middle schools, and one Grade 9-12 high school, with an approximate enrollment of about 6,000 students. The District Factor Group (DFG) is reported as B. DFGs are labeled from A (lowest) to J (highest) and are an indicator of the socioeconomic status of citizens in each district. According to the State of New Jersey School Report Card, the district is not currently making adequate yearly progress (AYP) and is classified as a district "in need of improvement in Year 3." The state of New Jersey categorizes "improvement status" by the following criteria.

Table 3

School Improvement Continuum Chart

NCLB/Title I
School Improvement Continuum Chart

Year	Status	Interventions for Title I Schools
Year 1	**Early Warning** – Did not make AYP for one year	None
Year 2	First year of **school in need of improvement** status. Did not make AYP for two consecutive years in the same content area.	Parent notification, public school choice (or supplemental educational services), school improvement plan, technical assistance from district.
Year 3	Second year of **school in need of improvement** status. Did not make AYP for three consecutive years in the same content area.	Parent notification, public school choice, supplemental educational services, school improvement plan, technical assistance from district.
Year 4	Third year of school in need of improvement status – **corrective action**. Did not make AYP for four consecutive years in the same content area.	Parent notification, public school choice, supplemental educational services, school improvement plan, technical assistance from district and state, corrective action, participation in CAPA.
Year 5	Fourth year of school in need of improvement status – **school restructuring** plan. Did not make AYP for five consecutive years in the same content area	Parent notification, public school choice, supplemental educational services, school improvement plan, technical assistance from district and state, development of restructuring plan (governance).
Year 6 and above	Fifth year of school in need of improvement status – **implementation of restructuring plan**. Did not make AYP for six consecutive years in the same content area.	Parent notification, public school choice, supplemental educational services, school improvement plan, technical assistance from district and state, implementation of restructuring plan.

New Jersey Department of Education (2009)

For the purposes of this study, the sample population was limited to the two Grade 6-8 middle schools in the city district. School A has approximately 710 students, 260 in Grade 6, 240 in Grade 7, and 210 in Grade 8. School B has approximately 700 students, 49 in Grade 5 (excluded for this study), 197 in Grade 6, 239 in Grade 7, and 216 in Grade 8. Ethnic/racial breakdown by grade level can be examined in the table below.

Table 4

District Demographics

	Asian		Black		Hispanic		American Indian / Alaskan		Multi-Racial		Hawaiian native/other Pacific Islander		White		All		Total
Grade	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	Total
P3A	0	0	2	3	1	3	0	0	0	0	0	0	1	2	**4**	**8**	**12**
	0.00	0.00	16.67	25.00	8.33	25.00	0.00	0.00	0.00	0.00	0.00	0.00	8.33	16.67	**33.33**	**66.67**	**100**
P3P	0	0	0	4	2	2	0	0	0	0	0	0	2	1	**4**	**7**	**11**
	0.00	0.00	0.00	36.36	18.18	18.18	0.00	0.00	0.00	0.00	0.00	0.00	18.18	9.09	**36.36**	**63.63**	**100**
PK3	0	0	7	2	5	3	0	1	3	0	0	0	2	0	**17**	**6**	**23**
	0.00	0.00	30.43	8.70	21.74	13.04	0.00	4.35	13.04	0.00	0.00	0.00	8.70	0.00	**73.91**	**26.09**	**100**
PK4	0	0	3	4	8	0	0	0	1	1	0	0	4	1	**16**	**6**	**22**
	0.00	0.00	13.64	18.18	36.36	0.00	0.00	0.00	4.55	4.55	0.00	0.00	18.18	4.55	**72.73**	**27.28**	**100**
P4F	3	2	38	36	38	48	0	1	6	4	0	0	44	46	**129**	**137**	**266**
	1.13	0.75	14.29	13.53	14.29	18.05	0.00	0.38	2.26	1.50	0.00	0.00	16.54	17.29	**48.51**	**51.5**	**100**
K	4	8	68	63	83	61	0	2	12	4	0	0	63	48	**230**	**186**	**416**
	0.96	1.92	16.35	15.14	19.95	14.66	0.00	0.48	2.88	0.96	0.00	0.00	15.14	11.54	**55.28**	**44.7**	**100**
01	6	4	63	67	70	73	0	0	7	10	0	1	65	73	**211**	**228**	**439**
	1.37	0.91	14.35	15.26	15.95	16.63	0.00	0.00	1.59	2.28	0.00	0.23	14.81	16.63	**48.07**	**51.94**	**100**
02	3	3	73	69	82	68	0	0	14	12	1	0	60	59	**233**	**211**	**444**
	0.68	0.68	16.44	15.54	18.47	15.32	0.00	0.00	3.15	2.70	0.23	0.00	13.51	13.29	**52.48**	**47.53**	**100**
03	6	2	68	67	66	61	0	0	9	5	0	0	52	51	**201**	**186**	**387**
	1.55	0.52	17.57	17.31	17.05	15.76	0.00	0.00	2.33	1.29	0.00	0.00	13.44	13.18	**51.94**	**48.06**	**100**
04	6	5	91	67	76	69	1	0	11	2	0	1	58	40	**243**	**184**	**427**
	1.41	1.17	21.31	15.69	17.80	16.16	0.23	0.00	2.58	0.47	0.00	0.23	13.58	9.37	**56.91**	**43.09**	**100**
05	4	5	74	83	82	66	0	0	4	2	0	0	49	39	**213**	**195**	**408**
	0.98	1.23	18.14	20.34	20.10	16.18	0.00	0.00	0.98	0.49	0.00	0.00	12.01	9.56	**52.21**	**47.8**	**100**
06	1	2	84	95	77	63	1	0	8	6	0	0	66	65	**237**	**231**	**468**
	0.21	0.43	17.95	20.30	16.45	13.46	0.21	0.00	1.71	1.28	0.00	0.00	14.10	13.89	**50.63**	**49.36**	**100**
07	11	5	86	88	82	75	0	0	6	5	1	0	66	59	**252**	**232**	**484**
	2.27	1.03	17.77	18.18	16.94	15.50	0.00	0.00	1.24	1.03	0.21	0.00	13.64	12.19	**52.07**	**47.93**	**100**
08	4	6	83	83	71	74	0	1	7	3	0	0	54	48	**219**	**215**	**434**
	0.92	1.38	19.12	19.12	16.36	17.05	0.00	0.23	1.61	0.69	0.00	0.00	12.44	11.06	**50.45**	**49.53**	**100**

Each school also has 100% of its staff deemed “highly qualified” in the content area in which he or she teaches as mandated by NCLB requirements. In other words, a teacher teaching Grade 6-8 mathematics holds either a K-12 or K-8 content area certification in mathematics and a teacher teaching Grade 6-8 language arts holds either a K-12 or K-8 content area certification in English/language arts. Additionally, School A and B have a combined count of 73.3% of school professionals holding a bachelors degree (BA/BS), while 26.7% of school professionals hold a masters degree (MA/MS). The classification of “school professionals” consists of teachers, administrators, guidance counselors, and speech therapists.

Participants who met the following criteria were chosen to participate in this study: (1) each student in the sample had valid overall and cluster scores in language arts/literacy and mathematics on the NJ ASK, (2) each student in the sample completed both previous grade levels in the same district and school (as indicated by obtaining two years of NJ ASK scores 2009-2010, 2010-2011, (3) each student in the sample was in grades 6-8 during the time of the study, (4) each student was considered a general education student and not deemed eligible for special education services.

Student participants were assigned to classrooms, both inclusive and non-inclusive, prior to the onset of this investigation by school district administration. While the researcher of this study was unable to control for class placement, pre-achievement scores could be used to get an overall achievement level for each group. Archival data were collected from student files. Achievement test scores were retrieved 2010-2011 via Genesis, the district student management software package. Non-disabled students were coded by grade level based on placement in

Comparison Group 1 (non-disabled students assigned to inclusive placements) or Comparison Group 2 (non-disabled students assigned to non-inclusive placements).

In this school district, an inclusive classroom is defined as an academic setting in which general education students and students in special education learn academics (for this study specifically math and language arts) in the same classroom environment. These students are taught in a classroom containing two teachers, one content area expert and one special education teacher. It is important to note that students in special education are not pulled from the general education setting for small group instruction at any time during "inclusion." Both the general education students and special education students are exposed to the same curriculum and materials as peers placed in non-inclusive settings.

Instrumentation NJ ASK 6, 7, and 8

The purpose of this study is to determine whether a significant relationship exists between the student/school variables discussed in the review of the literature and achievement in the district NJ ASK scores in Grades 6-8 in both language arts and mathematics. The NJ ASK scores determine the level of proficiency which is used as a standard measure throughout the state.

The New Jersey Assessment of Skills and Knowledge (NJ ASK) is the state test for students in Grades 3 through 8. It is designed to give schools information about how well students are achieving in the areas required by New Jersey's Core Curriculum Content Standards. The standards were adopted by the New Jersey State Board of Education after a public process that enlisted the help and advice of many educators, business representatives, and interested citizens. The standards are in the following areas:

1. Language Arts Literacy (Reading, Writing, Speaking, Listening, and Viewing)

2. Mathematics

3. Science

4. Visual and Performing Arts

5. Social Studies

6. Health and Physical Education

7. World Languages

8. Technological Literacy

9. Career Education and Consumer, Family, and Life Skills

In Grades 6-8 (as utilized by this study), students are tested in the areas of language arts, mathematics, and science in order to provide the state data on how well schools are preparing students to meet the goals of AYP (Annual Yearly Progress) and whether the students are being adequately prepared to pass the mandatory test given in eleventh grade to qualify for graduation from high school. According to the federal NCLB legislation, all students must be 100% proficient in both language arts and mathematics by the year 2014, thus proving that they are going to be both successful in school and in the competitive global economy.

According to the 2010 State NJ ASK Technical Report, Language Arts Literacy and Mathematics scores at Grades 3–8 and Science scores at grades 4 and 8 are reported as scale scores, with score ranges as follows:

- **Partially Proficient: 100-199**

  Students performing at the Partially Proficient level have limited recall, recognition, and application of basic facts and informational concepts.

- **Proficient: 200-249**

  Students performing at the Proficient level demonstrate recall, recognition, and application of facts and informational concepts.

- **Advanced Proficient: 250-300**

  Students performing at the Advanced Proficient level demonstrate the qualities outlined for Proficient performance. These students consistently demonstrate the ability to abstract relevant information, use multiple strategies and/or reasoning methods, and use various forms of representations to solve challenging problems. These students demonstrate an understanding of the reasonableness of their answers (NJDOE, 2009).

The scores of students who are included in the Partially Proficient level are considered to be below the state minimum of proficiency, and those students may be most in need of instructional support.

According to the New Jersey Department of Education standards for the 2010-2011 NJ ASK scores, school A and school B acquired the following test scores, putting them in the "in need of improvement" category as the percentage of students in the "proficient category" did not meet the state requirement for annual yearly progress. The following charts reflect the scores from School A and School B as well as a combined district percentage.

Table 5

2010-2011 NJ ASK 6: Language Arts/Literacy

School	Number Tested	Partial	Proficient	Advanced
A	207	53.6%	44.9%	1.4%
B	220	50%	49.1%	0.9%
District	427	51.8%	47.1%	1.2%

Table 6

2010-2011 *NJ ASK 6: Mathematics*

School	Number Tested	Partial	Proficient	Advanced
A	208	54.8%	39.9%	5.3%
B	218	39.4%	45%	15.6%
District	426	46.9%	42.5%	10.6%

Table 7

2010-2011 *NJ ASK 7: Language Arts/ Literacy*

School	Number Tested	Partial	Proficient	Advanced
A	238	54.8%	39.9%	5.3%
B	223	51.1%	43.5%	5.4%
District	461	50.5%	43.4%	6.1%

Table 8

2010-2011 NJ ASK 7: Mathematics

School	Number Tested	Partial	Proficient	Advanced
A	239	61.1%	32.6%	6.3%
B	223	53.8%	38.1%	8.1%
District	462	57.6%	35.3%	7.1%

Table 9

2010-2011 NJ ASK 8: Language Arts/Literacy

School	Number Tested	Partial	Proficient	Advanced
A	206	27.7%	66.5%	5.8%
B	227	16.7%	75.8%	7.5%
District	433	21.9%	71.4%	6.4%

Table 10

2010-2011 NJ ASK 8: Mathematics

School	Number Tested	Partial	Proficient	Advanced
A	207	46.9%	42%	11.1%
B	223	53.8%	38.1%	8.1%
District	432	39.1%	48.4%	12.5%

Reliability

Reliability refers to the consistency of the scores obtained throughout the research; in other words, how consistent they are for each individual from one administration of an

instrument to another and from one set of questions to another. If an instrument is reliable, then an individual who scores well initially should score within that same range during another administration. Perhaps the scores would not be identical; however, they would be close. "The New Jersey Department of Education is required by federal law to ensure that the instruments it uses to measure student achievement for school accountability provide reliable results" (NJDOE, 2009, p.116).

The NJ ASK assessment is built upon the principles of the Classical Test Theory (CTT). Members of the New Jersey Department of Education look at this approach in that "it builds on the notion of an ideal, error-free or true measurement score." Any observed measurement, such as a test score, X, is defined as a composite of true score, T, and its associated error: X = T + error" (NJDOE, 2009, p. 117). According to authors of the NJ ASK 2009 Technical Report,

> Estimating the size of the measurement error associated with the true score is the key to estimating reliability. Errors in measurement can result from any of a multitude of factors, including environmental factors (e.g., testing conditions) and examinee factors (e.g., fatigue, stress). CTT provides a means for this quantification of examinee inconsistency; i.e., measurement error (p. 117).

The report accounts for threats of reliability by making the statement that "When evaluating these results, it is important to recall that reliability is partially a function of test length; therefore, the reliability of a content area is likely to be greater than the reliability of a cluster simply because the content area has more items" (NJDOE, 2009, p. 12). Threats of reliability when using cluster scores instead of larger content areas cause concern as these scores

are commonly used to draw conclusions about teacher effectiveness, student achievement, and the success or failure of schools throughout the state.

Because this study is using the NJ ASK assessment as a means to draw some general conclusions about the influence of an inclusion/non-inclusion classroom on student achievement, "it is important to discuss the standard error of measure (SEM) in order to determine how the results rendered via the exam differ from a student's true score" (Tienken, 2008b, p. 37). Included in the NJ ASK Technical Report of 2009 is clarification on how the test creators deal with issues of SEM and reliability:

> Although the conceptualization of reliability and SEM is relatively straightforward, issues underlying the estimation of reliability are not. Reliability can be estimated via the correlation of scores on parallel forms or from test-retest data, or it can be estimated from a single test administration using any one of a variety of techniques (e.g., Brown, 191045; Cronbach, 195146; Kuder & Richardson, 193747). A very popular technique for estimating reliability from a single test administration is Cronbach's coefficient alpha. (NJDOE, 2009, p. 119).

Cronbach's α (alpha) is a coefficient of reliability. It is commonly used as a measure of the internal consistency or reliability of a psychometric test score for a sample of examinees. It was first named *alpha* by Lee Cronbach in 1951, as he had intended to continue with further coefficients. Cronbach's alpha will generally increase as the intercorrelations among test items increase and is thus known as an internal consistency estimate of reliability of test scores. Because intercorrelations among test items are maximized when all items measure the same construct, Cronbach's alpha is widely believed to indirectly indicate the degree to which a set of

items measures a single unidimensional latent construct (Fraenkel & Wallen, 2006). According to Gliem and Gliem (2003), reliability coefficients should be as close to 1 as possible, but commonly the following guidelines are used:

Table 11

Cronbach's Alpha

Cronbach's Alpha	Internal Consistency
$\alpha \geq .9$	Excellent
$.9 > \alpha \geq .8$	Good
$.8 > \alpha \geq .7$	Acceptable
$.7 > \alpha \geq .6$	Questionable
$.6 > \alpha \geq .5$	Poor
$.5 > \alpha$	Unacceptable

Fraenkel & Wallen (2007, p. 71)

The Standard Error of Measurement (SEM) estimates how repeated measures of a person on the same instrument are distributed around his or her true score. The larger the SEM, the lower the reliability of the test and the less precision there is in the scores.

The NJ ASK developers quantify student achievement on three different metrics: number correct or raw score, IRT scale, and performance score. Additionally, the New Jersey Department of Education (2009) states that "It is the responsibility of test developers to maximize reliability and minimize error by (1) identifying likely sources of error, (2) controlling

the conditions of error, (3) estimating the size of error and/or level of reliability, and (4) reporting the estimates by metric and unit of analysis" (p.75). The charts below indicate both the SEM and the alpha score for the seventh and eighth grade NJ ASK test as deemed the dependent variable of this study.

Table 12

Grade 7 Coefficient Alpha and Standard Error Measurement for Clusters

Table 8.1.6: Grade 7 Coefficient Alpha and Standard Error Measurement for Clusters

	Number of Items			Number of Possible Points	Alpha	SEM
	MC	CR/ECR	SCR			
LAL	**36**	**6**		**70**	**0.88**	**3.54**
Writing		2		18	0.50	1.94
Reading	36	4		52	0.88	2.85
Working with Text	21	0		21	0.78	1.96
Analyzing Text	15	4		31	0.80	2.04
Math	**35**	**3**	**8**	**52**	**0.91**	**3.27**
Number & Numerical Operations	8	1	2	13	0.72	1.55
Geometry & Measurement	8	1	2	13	0.73	1.71
Patterns & Algebra	11	0	2	13	0.74	1.50
Data Analysis, Probability, & Discrete Mathematics	8	1	2	13	0.71	1.63
Problem Solving	*15*	*3*	*6*	*30*	*0.87*	*2.59*

New Jersey Department of Education (2009)

Table 13

Grade 8 Coefficient Alpha and Standard Error Measurement for Clusters

Table 8.1.7: Grade 8 Coefficient Alpha and Standard Error Measurement for Clusters

	Number of Items			Number of Possible Points	Alpha	SEM
	MC	CR/ECR	SCR			
LAL	**36**	**6**		**70**	**0.89**	**3.33**
Writing		2		18	0.50	1.61
Reading	36	4		52	0.89	2.78
Working with Text	19	0		19	0.79	1.75
Analyzing Text	17	4		33	0.82	2.13
Math	**35**	**3**	**8**	**52**	**0.92**	**3.25**
Number & Numerical Operations	11	0	2	13	0.78	1.45
Geometry & Measurement	8	1	2	13	0.75	1.58
Patterns & Algebra	8	1	2	13	0.78	1.54
Data Analysis, Probability, & Discrete Mathematics	8	1	2	13	0.72	1.67
Problem Solving	*18*	*3*	*3*	*30*	*0.82*	*2.80*
Science	**48**	**6**		**54**	**0.87**	**3.55**
Life Science	21	0		21	0.75	2.02
Physical Science	13	1		16	0.61	2.02
Earth Science	14	1		17	0.74	1.90
Knowledge	*9*	*0*		*9*	*0.56*	*1.35*
Application	*39*	*2*		*45*	*0.86*	*3.12*

New Jersey Department of Education (2009)

When using standardized tests to make high stakes decisions, it is important to note the threats to reliability as this study utilizes local test scores in one urban school district. As indicated by the tables, the alpha score for Grade 7 is .88 (good range) with a SEM of 3.54 for language arts/literacy and .91 (excellent range) with a SEM of 3.27 for math. In Grade 8, the alpha score for language arts and literacy is .89 (good range) with a SEM of 3.33 and in math the alpha score is .92 (excellent) with a SEM of 3.25. Although these numbers outline a relatively "good" alpha score statewide, it is important to understand the limitations when they are translated into local school districts. Tienken (2008b) states:

> It is quite possible that test reliability will be much lower at the local level than for the state as a whole. The state does not provide district leaders with school-level reliability results; thus, leaders cannot gauge the actual amount of error they must factor into interpretation of test scores. In a district with very unreliable results the leaders might

unknowingly change a successful program or not recognize deficiencies in a program, but due to the size of the error in scores, the success or need for improvement might not be reflected in the results provided by the state. This may occur more frequently in districts with smaller testing populations because scores can fluctuate more from year to year because of changes in student characteristics (p. 39).

In summary, the reliability of the test scores is based on the statewide population and not the scores of the particular district. Although test scores must be reported to the public, there are many variations in which the information is presented. It has been found that some states report aggregated data with very specific details regarding alpha coefficients, while others simply report proficiency percentages; therefore, any recommendations or decisions rendered from this study will be made exercising caution, as the only statistics available are given via the state level.

NJ ASK Validity

Validity is the most important idea to consider when selecting an instrument (Fraenkel & Wallen, 2006). Each instrument must display evidence of careful test construction; adequate score reliability; appropriate test administration and scoring; accurate score scaling, equating, and standard setting; and careful attention to fairness for all examinees in order to demonstrate valid results. Fraenkel and Wallen define test validity as "the appropriateness, correctness, meaningfulness, and usefulness of the specific inferences researchers make based on the data that they collect" (p. 151). According to the authors of the NJ ASK Technical Report, creators of the NJ ASK consider the many forms of validity when designing the test, the steps in creating the program, and how the testing operating runs. "The validity of any assessment stems from the steps taken in planning it, the processes of developing the content of the tests, the processes of

consulting with stakeholders, the processes of communicating about the test to users, the processes of scoring and reporting, and the processes of data analysis. Each is an inherent part of validity" (p. 12).

Content validity refers to the content and format of the instrument. Baker and Linn (2002) suggest that "two questions are central in the evaluation of content aspects of validity. Is the definition of the content domain to be assessed adequate and appropriate? Does the test provide an adequate representation of the content domain the test is intended to measure?" (p. 6). The NJ ASK assessment was designed to measure a student's proficiency level according to the state-adopted New Jersey Core Curriculum Content Standards (NJCCCS). The NJCCCS were designed to provide each child across the state access to the same curriculum content and ensure that material was annually assessed by the NJ ASK in order to determine what the student understands and is able to do. Once the standards were placed in effect during 1996, updates to Administrative Code N.J.A.C. 6A:8 required districts to align all curriculum to the standards, requiring 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.

With that said, adequate representation of the content domains defined in the 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.

Threats to content validity arise within large-scale tests such as the NJ ASK. Large-scale state testing programs generally attempt to measure a wide array of knowledge and skills with a

relatively limited number of test questions (Tienken, 2008a). Tienken (2008a) continues to explain issues with validity in that "The NJ ASK assessment samples only a small part of a larger domain of content. It is difficult to determine from the information available publicly how well the items on the state tests represent the domains they attempt to assess" (p. 56).

Construct-based validity refers to the nature of the psychological construct or characteristic being measured by the instrument. In other words, it is the extent to which what was to be measured was actually measured. The NJ ASK was designed to assess student performance in several content areas using a variety of testing methods: multiple choice, open-ended responses, and essay. The NJ ASK 3-8 tests are scaled via raw score points, Item Response Theory (IRT), and performance standard level. Accordingly, "test developers actively promote the use of performance level results, reporting them annually on each content test at the student, school, district, and state levels. Individual student and average scale scores are also used but should play a secondary role, generally interpreted with reference to their distance from performance-score cut points. The testing manual instructs teachers and educational professionals to use NJ ASK performance scores as an indicator that an individual student performs at the Partially Proficient, Proficient, and Advanced Proficient level in a content area.

Data Collection

Permission was granted to me as the researcher to use all the requested sources of information by the district's superintendent of schools. Data were then gathered by the district testing coordinator and given via an Excel spreadsheet categorized by a student number. These numbers were assigned to an individual student in order to ensure that the data set remained anonymous and confidential. Each completed report included the following criteria: NJ ASK scores 2010 in math, language arts, and science (if available), NJASK scores 2011 in math,

language arts, and science (if available), eligibility for free and reduced lunch, student attendance record, gender, placement in an inclusion classroom or non-inclusion classroom, and whether a student was classified as special education. Because this study focuses on the academic achievement of general education students, students in special education were excluded from the analysis. Additionally, if one piece of data was missing or unavailable from a student's record, that student was also excluded for the purposes of this study.

Data Analysis

This study included convenience samples from two middle schools housing Grades 6-8 in an urban area in New Jersey where both academic subjects of language arts and mathematics were researched. All collected data were inputted in SPSS version 20 via computer. Through the use of multiple regression analysis, the predictor variables (student and school variables) were inputted as the independent variables, whereas the NJ ASK scores were inputted as the dependent variable(s), which is a scale level, dependent variable. According to Rubinfeld (1986),

> Multiple regression analysis is a statistical tool for understanding the relationship between two or more variables. Multiple regression involves a variable to be explained — called the dependent variable—and additional explanatory variables that are thought to produce or be associated with changes in the dependent variable. (p. 179)

Variables

The independent variables included economically disadvantaged, eligibility for free lunch, gender, race, having a 504 plan, special education status, current grade level, NJ ASK scores in language arts and math for the year 2009-2010, whether a student was placed in an inclusion setting for language and/or math, and the number of student absences. The dependent variables included NJ ASK scores in language arts and math for the year 2010-2011.

A variable was considered dichotomous when it took on only two values. In the case of all the variables except NJ ASK scores, the values could only be "yes" or "no." For example, "yes" the student is eligible for free lunch or "no" the student is not eligible for free lunch. The following recoding was used for the student variable of SES: 1 = eligible for free reduced lunch, O = not eligible. The coding for the SPSS data sheet is explained in the table below.

Table 14

SPSS Variable Coding

Variable	Measure	Coding
Economically Disadvantaged	Nominal	0= No 1= Yes
Free Lunch Eligible	Nominal	0= No 1= Yes
Gender	Nominal	0= Male 1=Female
Race	Nominal	0= White 1= Non-White
504 Plan	Nominal	0= No 1= Yes
Classified Special Education	Nominal	0= No 1= Yes
Grade Level	Ordinal	Grade 6, 7 or 8
LAL 2011	Scale	Scores Indicated
Proficiency LA11	Nominal	0= No 1= Yes
Math 2011	Scale	Scores Indicated
Proficiency MA11	Nominal	0= No 1= Yes
LAL 2010	Scale	Scores Indicated
Proficiency LA10	Nominal	0= No 1= Yes
Math 2010	Nominal	0= No 1= Yes
Proficiency MA10	Nominal	0= No 1= Yes

Variable	Measure	Coding
Inclusion Math	Nominal	0= No 1= Yes
Inclusion LA	Nominal	0= No 1= Yes
Days Absent	Scale	Number Indicated
School Attended	Nominal	0= SMS 1=MMS

As mentioned previously, the NJ ASK scores range from one of three levels, Partially Proficient (PP) 100-199, Proficient (P) 200-249, and Advanced Proficient (AP) 250-300. In order to examine the test data in numerous ways, another column of data was added for each row of scale test NJ ASK scores indicating whether a student's score was deemed Proficient (200 and above) or Non-proficient (199 and below). This column was coded in a nominal manner: 0 = Not Proficient, 1= Proficient. This procedure was done for each test score collected in both language arts and math content areas.

Chi-Square

Initially, I conducted a non-parametric chi-square to provide a quantitative measure of the relationship between two categorical variables: first, by determining what the distribution of observations (frequencies) would look like if no relationship existed and, second, by quantifying the extent to which the observed distribution differs from that determined in the first step. I used chi-squares to compare the special education sample populations between School A and B, the general education test scores in the areas of language arts and mathematics, and the racial make up of each school. After examining the chi squares (See Appendix A), I determined that the sample populations were not comparable enough to perform simultaneous regressions with the

entire data base. Some of the comparisons produced statistically significant differences between the sample populations of the two schools. For example, the chi square analysis provided a *p* value of $p<.001$ for race, meaning that there was a statistically significant difference between Schools A and B in terms of racial make up ($\chi2=51.981$, df=1,N=1280, $p<.001$). The same held true for economically disadvantaged ($\chi2=56.106$, df=1,N=1280, $p<.001$), and the special education population ($\chi2=4.229$, df=1,N=1280, $p<.05$). The test scores for 2011 NJ ASK language arts ($p=.221$) and 2011 NJASK mathematics ($p=.128$) were not found to be statistically significant.

These data provided a sound statistical rationale for performing a separate simultaneous regression analysis for each school per content area. The differences in schools violated a major regression assumption of normality (Leech, Barrett, & Morgan, 2001, p. 125). This violation compromises the estimation of coefficients and the calculation of confidence intervals. Sometimes the error distribution is skewed by the presence of a few large outliers. Since parameter estimation is based on the minimization of squared error, a few extreme observations can exert a disproportionate influence on parameter estimates. Calculation of confidence intervals and various significance tests for coefficients are all based on the assumptions of normally distributed errors. If the error distribution is significantly non-normal, confidence intervals may be too wide or too narrow. There were too many statistically significant differences in the school populations in order for them to be used in a combined sample; therefore, the data were analyzed in groups as follows: School A language arts, School A mathematics, School B language arts, School B mathematics.

Correlation Coefficients

Second, I ran a correlation analysis on the variables simultaneously. I removed any input variables that were clearly not good predictors of the output variable. It is essential that none of the input variables have a high correlation with (are good predictors of) other input variables. Additionally, after I ran the correlation analysis on the data, I removed any input variables that had a low correlation with the output variable. A correlation coefficient with an absolute value of less than 0.4 (between -0.4 and +0.4) between the output variable and an input variable indicates that the input variable is not a good predictor of the output (Witte & Witte, 2007). That input variable should be removed from the regression analysis.

Last, after looking at the correlation coefficients between the input and output variables, I examined the correlation coefficients between the input variables themselves. It is important not to use pairs of input variables that are good predictors of each other in a regression. This will cause a regression error known as collinearity or multicollinearity. One variable from any pair of highly correlated input variables should be removed prior to running the regression analysis. Variables can be considered highly correlated if the absolute value of their correlation coefficient is greater than 0.7 (greater than +0.7 or less than -0.7) (Witte & Witte, 2007).

Multiple Regression

I used simultaneous multiple regression as the first analysis strategy. In order to ensure that a proper analysis procedure was used for this study, I used the Enter method of the SPSS program (also known as simultaneous regression) where all appropriate student and school variables were entered at the same time. I ran two multiple regression analyses for each school, one for language arts and one for math and therefore generated a total of 4 model summaries. The analyses were done revealing the following information:

Explanation of Variance: Determining how much of the variance in the NJ ASK scores can be explained by student placement in an inclusion classroom.

Significance of the Regression Equation: Revealing if the regression equation is statistically (p value $\leq$.005)

Explanation of Coefficients: The examination of the Standardized Coefficient indicates the direction (positive or negative) and whether the variables have a significant influence on the NJASK scores. The Beta (b), t value, and the p value are identified. The p value determines significance. It is important to note that the closer the Beta (b) to 1, the stronger the influence of the predictor. The examination of the Beta (slope) for the predictor variable will indicate the influence as weak, moderate, or strong.

Conducting a multiple regression analysis provided valuable statistical information that led me to be able to make data driven interpretations and recommendations due to an improved methodology, to add to the limited body of empirical evidence currently in existence addressing the placement of general education students in an inclusive setting.

Hierarchical Regression

In addition to simultaneous regression, I used a hierarchical regression analysis. Hierarchical regression adds terms to the regression model in stages based on a sequential entering of each variable. At each stage, an additional variable or variables were added to the model and the change in R^2 was calculated. A hypothesis test was done to test whether the change in R^2 was significantly different from zero. The first table is a table of what variables

were entered or removed at the different stages (See table 15). For the purposes of this study, the inclusion variable was entered first, followed by the other student variables of economically disadvantaged, race, gender, and student attendance. Last, the variable of previous achievement was added to complete the block entry.

Table 15

Hierarchical Regression Model

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	INC Math[b]	.	Enter
2	Gender, Absent, Economically Disadvantaged, Race[b]	.	Enter
3	MAth 2010[b]	.	Enter

a. Dependent Variable: Math 2011

b. All requested variables entered.

After the variables were entered, the remainder of the tables displayed the following: the second table is a summary of the results of the different models, the third table presents the ANOVA significance table for the three models, and the fourth table contains the regression weights and significance levels for each model. As before, the *p,* or significance column, is a hypothesis test of the significance of that variable, given all the other variables at that stage have been entered into the model. The hierarchical regression analysis verified the results produced by the simultaneous regression and provided an opportunity to show the change in R^2.

Consequently, because this is a study that utilizes both a simultaneous and multiple regression analysis with more than four predictors, an examination of the correlation tables revealed a significant relationship between a few of the variables entered, the partial correlation coefficient was reported for all significant variables. A partial correlation coefficient represents the correlation between the criterion and a predictor variable. After removing variance that the predictor has in common with other predictors, the partial coefficient expresses the correlation between the predictor and the unaltered criterion. As a result, a closer approximation of the amount of variance explained by that particular predictor variable on the outcome variable can be determined. The partial correlation coefficient is reported as a range along with the Beta when the percentage of variance each variable accounts for is stated (Leech, Barrett, & Morgan, 2011, p. 125).

Factorial ANOVA

Last, additional analysis was required based on the statistical findings of the hierarchical and simultaneous multiple regression models in both Schools A and B, which indicated the potential for a difference in inclusion between the two schools. Furthermore, a factorial ANOVA was used to render further suggestions regarding the results of this study and the interaction between inclusion and school code on the academic achievement of general education students.

Factorial ANOVA measures whether a combination of independent variables predict the value of a dependent variable. As stated by Leech, Barrett, and Morgan (2011), this statistic is used when there are two different independent variables, each of which classifies or labels participants with respect to a particular characteristic, with each participant being labeled by a particular level of each of the independent variables (p. 223). ANOVA provides the opportunity

to examine power--with the same sample size and effect size. Also, a factorial ANOVA analysis can detect interactions or when the effects of one independent variable differ according to levels of another independent variable (Leech, Barrett & Morgan, 2011, p. 125). The profile plots of cell means help to visualize the nature of a significant interaction if and when one exists.

Chapter IV

ANALYSIS OF THE DATA

Introduction

My purpose for this study was to determine whether placement in an inclusive setting influences the academic achievement of general education students on the Language Arts Literacy and Mathematics section of the New Jersey Assessment of Skills and Knowledge (NJ ASK) 6,7, and 8. Additionally, I aimed to examine specific models, including the independent variables of inclusive setting, non-inclusive setting, gender, race, student attendance, and eligibility for free and reduced lunch that, paired with placement in an inclusive/non-inclusive setting, may influence the dependent variable of student achievement on the NJ ASK 6-8. As emphasis continues to be placed on the outcomes of the NJ ASK scores to determine school accountability, district administrators must consider the needs of all students. With that said, I designed this study to produce research-based evidence to continue the development of policy and create the opportunity for educated, data-driven decisions that benefit all students in inclusive settings. As a result, school administrators and teachers will be able to choose an instructional program that meets individual student needs of both disabled and non-disabled students that will maximize both learning and student achievement.

Research Questions

The individual SPSS outputs will be used to answer the following research questions:

Research Question 1: What is the influence of placement in the inclusive setting on the

performance of non-disabled students in the area of mathematics as measured by the NJ ASK when controlling for student mutable variables in Grades 6, 7, and 8?

Research Question 2: What is the influence of placement in the inclusive setting on the performance of non-disabled students in Grades 6, 7, and 8 in the area of language arts and literacy when controlling for student mutable variables as measured by the NJ ASK?

Research Question 3: How well does placement in an inclusion classroom, student gender, attendance in class, race, and SES predict and/or influence student overall academic performance as measured by the NJ ASK 6, 7, and 8?

Null Hypotheses:

H_o^1: Placement in an inclusive setting has no significant influence on the performance of non-disabled students in Grades 6-8 in the area of mathematics as measured by the NJ ASK.

H_o^2: Placement in an inclusive setting has no significant influence on the performance of non-disabled students in Grades 6-8 in the area of language arts and literacy as measured by the NJ ASK.

H_o^3: Placement in an inclusion classroom, student gender, attendance in class, race, and SES has no significant influence on the performance of non-disabled students in the areas of language arts and mathematics as measured by the NJ ASK.

Results

Language Arts, School A

Multiple regression analysis was performed to determine the amount of variability in the dependent or outcome variable NJ ASK Language Arts and Literacy 2011 (LAL 2011) that could be explained by the student demographic variables of (a) placement in an inclusion classroom for

language arts (INC LA), (b) eligibility for free lunch (free lunch eligible), (c) gender, (d) race, and (e) number of days absent (absent). Table 16 provides the descriptive statistics for both the outcome variable and all the predictor variables used in the model.

Table 16

Descriptive Statistics for Language Arts School A

Descriptive Statistics

	Mean	Std. Deviation	N
LAL 2011	200.55	22.578	486
INC LA	0.34	0.473	486
LAL 2010	198.35	23.369	486
Free Lunch Eligible	0.48	0.500	486
Gender	0.52	0.500	486
Race	0.81	0.389	486
Absent	4.60	5.026	486

First, a correlation table (Appendix B) was examined in order to identify relationships between a dependent variable and one or more independent variables. Correlation and regression analysis are related in the sense that both deal with relationships among variables. The correlation coefficient is a measure of linear association between two variables. Values of the correlation coefficient are always between -1 and +1. A correlation coefficient of +1 indicates that two variables are perfectly related in a positive linear sense, a correlation coefficient of -1 indicates that two variables are perfectly related in a negative linear sense, and a correlation coefficient of 0 indicates that there is no linear relationship between the two variables.

With that said, the Pearson Correlation table reveals that there is a strong negative relationship between the predictor variable of INC LA and the dependent variable of LAL 2011 (Pearson r= -.580) that was found to be statistically significant (*p<.001).* Also, there is a strong relationship between the predictor variable LAL 2010 and the dependent variable LAL 2011 (Pearson r= .714) which was found to be statistically significant (*p<.001).* There is expected to be a high correlation for these two variables considering that they are relatively the same exam given in two different school years 2010 and 2011. Finally, there is a negative moderate relationship between the predictor variable of LAL 2010 and predictor variable INC LA (Pearson r=-.433) which was found to be statistically significant (*p<.001).*

These findings indicate a need to report the partial correlation coefficients for each of these variables because one variable could on occasion act as a suppressor to the other variable. Partial correlations are important when there are four or more predictor variables in a simultaneous regression model and/or when you have a strong correlation between two or more predictor variables in any regression model that is run even though the tolerances might be within acceptable ranges. The partial correlation coefficient partitions out the influence of highly related predictor variables on one another. In other words, there is a closer approximation of the amount of variance explained by that particular predictor variable on the outcome variable (Leech, Barrett & Morgan, 2011). For each analysis, the Beta is reported as well as the partial correlation coefficient, as it displays the range in percentage of overall variance accounted for in each model on the dependent variable of NJ ASK LAL 2011 test scores. Using the linear regression analysis function on SPSS, a statistically significant model emerged (See Tables 17 through 20).

Table 17

Variables Entered/Removed

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Absent, Gender, INC LA, Race, Free Lunch Eligible, LAL 2010[b]	.	Enter

a. Dependent Variable: LAL 2011

b. All requested variables entered.

Table 18

Model Summary Language Arts School A

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.779[a]	.606	.601	14.255	.606	122.946	6	479	.000

a. Predictors: (Constant), Absent, Gender, INC LA, Race, Free Lunch Eligible, LAL 2010

The adjusted R^2 for this model indicates that 60.1% of the variance in student performance on the NJ ASK 6-8 in language arts can be explained by student absence, gender, placement in an inclusion classroom for language arts, race, free lunch eligibility, and the 2010 NJ ASK scores in Language Arts and Literacy. Additionally, this regression model (Table 18) is statistically significant with F=122.946, df=6,479, p <.001 (See Table 19).

Table 19

ANOVA Table for Language Arts School A School

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	149899.070	6	24983.178	122.946	.000[b]
	Residual	97335.039	479	203.205		
	Total	247234.109	485			

a. Dependent Variable: LAL 2011

b. Predictors: (Constant), Absent, Gender, INC LA, Race, Free Lunch Eligible, LAL 2010

Examination of the standardized coefficient (Table 20) indicates that there are two statistically significant predictors, INC LA and LAL 2010. Multicollinearity was not an issue because the tolerance values for each variable in the model was greater than .399 (<1-R^2) and all reported variance inflation factors (VIFs) were close to 1.

Table 20

Coefficients Table for Language Arts School A

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	101.349	6.843		14.811	.000					
	Gender	2.552	1.305	.057	1.956	.051	.120	.089	.056	.984	1.016
	Absent	-.106	.130	-.024	-.815	.415	-.031	-.037	-.023	.979	1.021
	Race	-2.458	1.725	-.042	-1.425	.155	-.167	-.065	-.041	.931	1.074
	Free Lunch Eligible	-1.264	1.338	-.028	-.944	.345	-.132	-.043	-.027	.935	1.069
	LAL 2010	.536	.031	.555	17.162	.000	.714	.617	.492	.787	1.271
	INC LA	-15.794	1.520	-.331	-10.390	.000	-.580	-.429	-.298	.809	1.235

a. Dependent Variable: LAL 2011

The predictor of placement in an inclusion class for language arts (INC LA) has a negative and significant influence on general education students' performance on the NJ ASK assessment in the area of language arts (β= -.331, t=-10.390, p<.001). As per the dummy coding criteria, the negative Beta -.331 indicates that general education students who are not in inclusion classrooms for language arts outperform general education students who are placed in inclusion settings for language arts. It is also important to note that the closer the Beta to 1, the stronger the influence of the predictor. The Beta .331 indicates that although it is a predictor of student performance on the NJ ASK in language arts, it is a moderate predictor, contributing between 11% to 18% of the overall variance in LAL 2011 performance for this model.

The predictor of LAL 2010 had a positive and significant influence on general education students' performance on the NJ ASK assessment in the area of language arts (β=.555, t=17.163, p<.001). The higher the previous year's 2010 NJASK score in the area of language arts, the better the students performed on the NJ ASK test during the 2011 school year. Again, the closer the Beta is to 1, the stronger the influence of the predictor. The Beta .555 indicates that it is a moderately strong predictor of student performance on the NJ ASK 2011 in language arts, contributing between 31% to 38% of the overall variance in LAL 2011 performance for this model. Additionally, to compare the influence of the two significant variables LAL 2010 and INC LA, the Beta .331 and .555 from each variable can be examined by dividing .555 by .331. With that said, LAL 2010 has a 1.67 times greater influence on student performance in language arts than placement in an inclusive language arts classroom.

The predictor of gender has a positive and arguably significant influence on student performance on the NJ ASK in language arts (β=.057, t=1.956, p<.051). Due to the positive

Beta .057 and the dummy coding of 0= males and 1=females, these results indicate that females are marginally outperforming males on the NJ ASK in language arts. The closer the Beta is to 1, the stronger the influence of the predictor. The Beta of .057 indicates that gender is an extremely weak predictor of student performance on the NJ ASK in language arts, contributing only between .3% to .8% of the overall variance in LAL 2011 performance for this model.

The independent variables of free lunch eligible, race, and absences were not found to be statistically significant predictors of performance on the 2011 NJ ASK LAL assessment as the *p* values for these variables were greater than .05.

Given the results produced by the simultaneous linear regression models, additional clarity was required in order to more thoroughly examine the R^2 change between the variables. Hierarchical multiple regression is a variant of the basic multiple regression procedure that allows the researcher to specify a fixed order of entry for variables in order to control for the effects of covariates or to test the effects of certain predictors independent of the influence of others (Leech, Barrett, & Morgan, 2011). Because the outputs indicated that previous years test scores were the strongest predictor of student achievement on the NJASK 2011 in language arts, the hierarchical regression analysis allowed the variables to be entered in various models as shown below.

Table 21

Hierarchical Regression Block Inputs, School A Language Arts

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	INC LA[b]	.	Enter
2	Absent, Gender, Economically Disadvantaged, Race[b]	.	Enter
3	LAL 2010[b]	.	Enter

a. Dependent Variable: LAL 2011

b. All requested variables entered.

As displayed in Table 21, variables were entered into the regression models as per the following blocks: Model 1, Inclusion LA; Model 2, Inclusion LA, Absent, Gender, Economically Disadvantaged, Race; Model 3, Inclusion LA, Absent, Gender, Economically Disadvantaged, Race, LAL 2010 (previous year's scores). By entering the variables in block models, any observed effect on LAL 2011 scores can then be said to be "independent of" the effects of the variables previously controlled for.

Table 22

Model Summary Hierarchical Regression, School A Language Arts

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.580[a]	.337	.335	18.406	.337	245.749	1	484	.000
2	.604[b]	.365	.359	18.082	.028	5.376	4	480	.000
3	.779[c]	.607	.602	14.236	.242	295.377	1	479	.000

a. Predictors (Constant), INC LA

b. Predictors: (Constant), INC LA, Absent, Gender, Economically Disadvantaged, Race

c. Predictors (Constant), INC LA, Absent, Gender, Economically Disadvantaged, Race, LAL 2010

It is also essential to examine the R squared change, which determines the percentage of variability in the dependent variable that can be accounted for by all the predictors together. The change in R squared is a way to evaluate how much predictive power was added to the model by the addition of another variable. The hierarchical regression method will also determine if that R squared change is significant as each block model is introduced. In examination of the model summary, for Model 1, the R squared change began at .337. The analysis of the adjusted R squared indicates that INC LA accounts for approximately 33.5% of the variability, which is deemed statistically significant ($p<.001$). When the additional variables are entered in Model 2, the R squared change is minimal, .028, and the percentage of variability (adjusted R squared) accounted for changes from .335 (33.5%) to .359 (35.9%) or 2.4%. Although a significant change, this translates into Model 2 explaining a very small proportion of the variance. However, when Model 3 is added, which adds the previous year's test scores, there is an R squared change of .242 (24.2%), and the adjusted R squared is .607, meaning that 60.7% of the variance is now accounted for when all of the variables are entered into the regression.

Table 23

Hierarchical Regression ANOVA Table for School A, Language Arts

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	83258.169	1	83258.169	245.749	.000[b]
	Residual	163975.940	484	338.793		
	Total	247234.109	485			
2	Regression	90289.796	5	18057.959	55.229	.000[c]
	Residual	156944.313	480	326.967		
	Total	247234.109	485			
3	Regression	150154.314	6	25025.719	123.479	.000[d]
	Residual	97079.795	479	202.672		
	Total	247234.109	485			

a. Dependent Variable: LAL 2011

b. Predictors: (Constant), INC LA

c. Predictors: (Constant), INC LA, Absent, Gender, Economically Disadvantaged, Race

d. Predictors: (Constant), INC LA, Absent, Gender, Economically Disadvantaged, Race, LAL 2010

This table confirms the previous results as well as indicating that each model in and of itself is significant. The independent variables entered in Models 1, 2, and 3 predicted scores on the NJ ASK 2011 LAL test in language arts to a statistically significant degree ($p<.001$) in school A (Model 1: $F=245.749$, $df= 1, 484$, $p<.001$; Model 2: $F= 55.229$, $df= 5, 480$, $p<.001$; Model 3: $F= 123.479$, $df= 6, 479$, $p<.001$).

Table 24

Hierarchical Regression Coefficients Table for School A Language Arts

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	209.894	1.026		204.626	.000					
	INC LA	-27.681	1.766	-.580	-15.676	.000	-.580	-.580	-.580	1.000	1.000
2	(Constant)	215.434	2.356		91.456	.000					
	INC LA	-26.486	1.758	-.555	-15.070	.000	-.580	-.567	-.548	.974	1.027
	Economically Disadvantaged	-3.980	1.797	-.084	-2.214	.027	-.178	-.101	-.081	.923	1.083
	Gender	4.022	1.652	.089	2.435	.015	.120	.110	.089	.988	1.012
	Race	-5.545	2.201	-.096	-2.519	.012	-.167	-.114	-.092	.920	1.087
	Absent	-.197	.164	-.044	-1.205	.229	-.031	-.055	-.044	.997	1.003
3	(Constant)	102.032	6.854		14.886	.000					
	INC LA	-15.660	1.520	-.328	-10.300	.000	-.580	-.426	-.295	.807	1.239
	Economically Disadvantaged	-2.083	1.419	-.044	-1.467	.143	-.178	-.067	-.042	.918	1.090
	Gender	2.521	1.303	.056	1.935	.054	.120	.088	.055	.984	1.016
	Race	-2.143	1.744	-.037	-1.229	.220	-.167	-.056	-.035	.908	1.101
	Absent	-.119	.129	-.026	-.920	.358	-.031	-.042	-.026	.996	1.004
	LAL 2010	.535	.031	.554	17.187	.000	.714	.618	.492	.789	1.267

a. Dependent Variable: LAL 2011

The coefficients table provides a detailed analysis of the strength of each individual predictor. In Model 1, the predictor of INC LA is statistically significant, $p<.001$ with t= -15.676 and a β= -.580. Independently, INC LA is a moderate predictor of student performance on the NJ ASK assessment in Language Arts because the closer the Beta to 1, the stronger the strength of the predictor. Additionally, because the Beta is negative, this indicates that non-disabled students that are not in inclusion are doing better than non-disabled students assigned to an inclusion classroom contributing 33% to the overall variance in NJASK LAL 2011 performance for this model.

When the variables of economically disadvantaged, gender, race, and student absences were added to INC LA, the strength of the variable INC LA decreased by only .03. (-.580 vs. -.555), meaning that these variables have a minimal effect on the strength of INC LA, as β= -.555 is still a moderate predictor of student performance on the NJASK 2011 in Language Arts,

contributing between 30.8% to 32% of the overall variance in LAL 2011 performance for this model. It is important to note that although INC LA was the strongest predictor in Model 2, other predictors were very weak predictors, although they were found to be statistically significant.

Economically disadvantaged was found to be a statistically significant variable (β= -.084, t= -2.214, *p*<.027), contributing between .7% to 1% of the overall variance in LAL 2011 performance for this model. It can be said that there is a slight but minimal relationship identified between economically disadvantaged students and the LAL 2011 scores but that relationship could be due to chance or error since it is not significant. With a Beta -.084, economically disadvantaged is a weak predictor of student performance on the NJ ASK 2011 in language arts.

Also, race was found to be a statistically significant predictor of student performance on the NJ ASK in language arts (β=-.096, t=-2.519, *p*<.012), contributing between .9% to 1.2% of the overall variance in LAL 2011 performance in this model. The Beta for the predictor race is -096, making it a very weak predictor of student performance on the NJ ASK 2011 in language arts.

Student absences and gender were not found to be statistically significant predictors of student performance on the NJASK in language arts (*p*>.05).

In Model 3, when LA 2010 (previous year's scores) were added to the existing variables, a lot of changes occurred. First, economically disadvantaged and race were no longer statistically significant variables (p<.143 & p<.220). LA 2010 was found to be a moderately statistically significant predictor (β=.554, t=17.187, p<.001) of student performance on the

NJASK 2011 LAL, contributing between 31% to 38% of the overall variance in LAL 2011 performance for this model, meaning that non-disabled students who did well on previous year's NJ ASK, also did well on the current year's test in language arts. Although the Beta of INC LA dropped from -.555 to -.328, it is still a statistically significant variable; however, its strength has decreased, becoming a weaker predictor of student performance on the NJASK in language arts (β= -.328, t=-10.300, p<.001), now contributing 10% of the overall variance in NJ ASK LAL 2011 performance for this model. It is important to note the value when the partial correlation is examined because in this case it is partitioning out the influence of the other variables in this model, specifically the influence of LAL 2011. The Beta for INC LA indicates that it accounts for only 10% of the variance in NJ ASK LAL 2011 scores; however, the partial correlation tells us that INC LA accounts for between 10% to 18% of the variance in NJ ASK LAL 2011 which indicates that the LA 2010 test scores have an influence over INC LA and are somewhat suppressing the strength of the INC LA variable.

Gender is not statistically a significant variable, but a notable variable, and is a weak predictor of student performance on the NJ ASK 2011 in language arts contributing .3% to 7% of the overall variance in LAL 2011 for this model (β=.056, t= 1.935, p<.054). It can be said that both the simultaneous regression as well as the hierarchical regression both make one fact clear, previous year test scores (LA 2010) and placement in an inclusion setting (INC LA) are the strongest predictors of student performance on the NJASK 2011 in language arts and explain the largest proportion of the variance for School A in language arts.

Mathematics, School A

A second multiple regression analysis was performed to determine the amount of variability in the dependent or outcome variable NJ ASK Math 2011 (Math 2011) that could be explained by the student demographic variables of (a) placement in an inclusion classroom for math (INC Math), (b) eligibility for free lunch (free lunch eligible), (c) gender, (d) race, and (e) number of days absent (absent). Table 25 provides the descriptive statistics for both the outcome variable and all the predictor variables used in the model.

First, a correlation table (Appendix B) was examined in order to identify relationships between a dependent variable and one or more independent variables. Correlation and regression analysis are related in the sense that both deal with relationships among variables.

The Pearson Correlation table reveals that there is a strong negative relationship between the predictor variable of INC Math and the dependent variable of Math 2011 (Pearson $r = -.681$) that was found to be statistically significant ($p < .001$). Also, there is a strong relationship between the predictor variable Math 2010 and the dependent variable Math 2011 (Pearson $r = .797$), which was found to be statistically significant ($p < .001$). There is expected to be a high correlation for these two variables considering that they are relatively the same exam given in two different school years, 2010 and 2011. Finally, there is a negative moderate relationship between the predictor variable of Math 2010 and predictor variable INC Math (Pearson $r = -.595$), which was found to be statistically significant ($p < .001$).

These findings indicate a need to report the partial correlation coefficients for each of these variables because one variable could on occasion act as a suppressor to the other variable when a simultaneous regression with 4 or more predictors is run. The partial correlation

coefficient calculation partitions out the influence of highly related predictor variables on one another. In other words, there is a closer approximation of the amount of variance explained by that particular predictor variable on the outcome variable (Leech, Barrett, & Morgan, 2008). For each analysis, the Beta is reported as well as the partial correlation coefficient, as it displays the range in percentage of overall variance accounted for in each model on the dependent variable of NJ ASK Math 2011 test scores. Using the linear regression analysis function on SPSS, a statistically significant model emerged (See Tables 25 through 28).

Table 25

Descriptive Statistics for Mathematics, School A

Descriptive Statistics

	Mean	Std. Deviation	N
Math 2011	211.12	33.187	486
Absent	4.60	5.026	486
Free Lunch Eligible	.48	.500	486
Gender	.52	.500	486
Race	.81	.389	486
INC Math	.32	.468	486
MAth 2010	212.74	34.858	486

Table 26

Variables Entered/Removed

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	MAth 2010, Absent, Gender, Free Lunch Eligible, Race, INC Math[b]	.	Enter

a. Dependent Variable: Math 2011

b. All requested variables entered.

Table 27

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.840[a]	.705	.701	18.136	.705	190.850	6	479	.000

a. Predictors: (Constant), MAth 2010, Absent, Gender, Free Lunch Eligible, Race, INC Math

The adjusted R^2 for this model indicates that 70.1% of the variance in student performance on the NJ ASK 6-8 in Mathematics can be explained by student absence, gender, placement in an inclusion classroom for language arts, race, free lunch eligibility, and the 2010 NJASK scores in Language Arts and Literacy. Additionally, this regression model (Table 26) is statistically significant with $F=190.850$, $df=6,479$, $p<.001$.

Table 28

ANOVA Mathematics, School A

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	376625.061	6	62770.843	190.850	.000[b]
	Residual	157543.532	479	328.901		
	Total	534168.593	485			

a. Dependent Variable: Math 2011

b. Predictors: (Constant), MAth 2010, Absent, Gender, Free Lunch Eligible, Race, INC Math

Examination of the standardized coefficients (Table 29) indicates that there are three statistically significant predictors, INC Math, gender and Math 2010. Multicollinearity was not

an issue because the tolerance values for each variable in the model was greater than .399 (< 1 – R^2) and all reported VIFs were close to 1.

Table 29

Coefficients Table for Mathematics, School A

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	100.968	7.719		13.081	.000				.	
	Gender	-3.534	1.660	-.053	-2.129	.034	-.132	-.097	-.053	.984	1.016
	Absent	-.161	.166	-.024	-.967	.334	-.094	-.044	-.024	.972	1.029
	Race	-1.163	2.218	-.014	-.524	.600	-.181	-.024	-.013	.912	1.096
	Free Lunch Eligible	1.034	1.700	.016	.608	.543	-.098	.028	.015	.938	1.066
	MAth 2010	.567	.030	.595	18.717	.000	.797	.650	.464	.608	1.644
	INC Math	-22.934	2.191	-.323	-10.466	.000	-.681	-.431	-.260	.644	1.552

a. Dependent Variable Math 2011

The predictor of placement in an inclusion class for mathematics (INC MATH) has a negative and significant influence on general education students' performance on the NJ ASK 2011 assessment in the area of mathematics (β= -.323, t=-10.466, p<.001). As per the dummy coding criteria, the negative Beta -.323 indicates that general education students who are not in inclusion classrooms for mathematics outperform general education students who are placed in inclusion settings for mathematics. It is also important to note that the closer the Beta to 1, the stronger the influence of the predictor. The Beta of .323 indicates that although it is a predictor of student performance on the NJ ASK 2011 in mathematics, it is a moderate predictor contributing between 10.4%-18.5% of the overall variance in NJ ASK 2011 math performance for this model.

The predictor of Math 2010 had a positive and significant influence on general education students' performance on the NJ ASK assessment in the area of mathematics (β=.595, t=18.717,

$p<.001$). The higher the previous year's 2010 NJ ASK score in the area of mathematics, the better the students did on the NJ ASK test during the 2011 school year. Again, the closer the Beta is to 1, the stronger the influence of the predictor. The Beta .595 indicates that it is a moderate predictor of student performance on the NJ ASK 2011 in mathematics, contributing between 35% to 42% of the overall variance in Math 2011 performance of this model. Additionally, Math 2010 has a 1.84 times greater influence on student performance in mathematics than placement in an inclusive mathematics classroom, which is calculated by dividing .595 by .323, the Beta for INC Math.

The predictor of gender has a positive and significant influence on student performance on the NJ ASK 2011 in mathematics (β=-.053, t=-2.129, $p<.034$). Due to the negative Beta -.053 and the dummy coding of 0= males and 1=females, these results indicate that males are marginally outperforming females on the NJ ASK 2011 in mathematics. The closer the Beta is to 1, the stronger the influence of the predictor. The Beta of .053 indicates that gender is an extremely weak predictor of student performance on the NJASK in mathematics contributing only between .2% to .9% of the overall variance in Math 2011 performance for this model.

The independent variables of free lunch eligible, race, and absences were not statistically significant predictors of performance on the 2011 NJ ASK assessment as the *p* value was greater than .05.

Given the results produced by the simultaneous linear regression models, additional clarity was required in order to more thoroughly examine the R squared change between the variables. Hierarchical multiple regression is a variant of the basic multiple regression procedure that allows the researcher to specify a fixed order of entry for variables in order to control for the

effects of covariates or to test the effects of certain predictors independent of the influence of others (Leech, Barrett, & Morgan, 2011). Because the outputs indicated that the previous year's test scores were the strongest predictor of student achievement on the NJ ASK 2011 in mathematics, the hierarchical regression analysis allowed the variables to be entered into various models as shown below.

Table 30

Hierarchical Regression Variables Entered, Mathematics School A

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	INC Math[b]	.	Enter
2	Gender, Absent, Economically Disadvantaged, Race[b]	.	Enter
3	MAth 2010[b]	.	Enter

a. Dependent Variable: Math 2011

b. All requested variables entered.

As displayed in Table 30, variables were entered into the regression models as per the following blocks: Model 1, Inclusion Math; Model 2, Inclusion Math, Absent, Gender, Economically Disadvantaged, Race; Model 3, Inclusion Math, Absent, Gender, Economically Disadvantaged, Race, Math 2010 (previous year's scores). By entering the variables in block models, any observed effect on Math 2011 scores can then be said to be "independent of" the effects of the variables previously controlled for.

Table 31

Model Summary Hierarchical Regression, School A Mathematics

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.681[a]	.464	.463	24.314	.464	419.563	1	484	.000
2	.700[b]	.490	.484	23.828	.025	5.991	4	480	.000
3	.840[c]	.705	.701	18.141	.215	349.143	1	479	.000

a. Predictors: (Constant), INC Math

b. Predictors (Constant), INC Math, Gender, Absent, Economically Disadvantaged, Race

c. Predictors: (Constant), INC Math, Gender, Absent, Economically Disadvantaged, Race, MAth 2010

It is also essential to examine the R squared change which examines the percentage of variability in the dependent variable that can be accounted for by all the predictors together. The change in R squared is a way to evaluate how much predictive power was added to the model by the addition of another variable. The hierarchical regression method will also determine if the R squared change is significant as each block model is introduced. In examination of the model summary, for Model 1, the R squared change began at .464. Analysis of the adjusted R squared indicates that INC Math accounts for approximately 46.3% of the variability, which is deemed statistically significant ($p<.001$). When the additional variables are entered in Model 2, the R squared change is minimal, .025 (2.5%), and the percentage of variability (adjusted R squared) accounted for changes from .463 (46.3%) to .484 (48.4%), or 2.1%. Although Model 2 only contributed a small proportion of the variance when compared to Model 1, Model 2 is a stronger predictor, accounting for more percentage of variance explained. However, when Model 3 is added, which adds the previous year's test scores, there is an R squared change of .215 (21.5%), and the adjusted R squared is .705, meaning that 70.5% of the variance is now accounted for when all of the variables are entered into the regression.

Table 32

Hierarchical Regression ANOVA Table for School A, Mathematics

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	248037.414	1	248037.414	419.563	.000[b]
	Residual	286131.179	484	591.180		
	Total	534168.593	485			
2	Regression	261643.126	5	52328.625	92.167	.000[c]
	Residual	272525.467	480	567.761		
	Total	534168.593	485			
3	Regression	376539.258	6	62756.543	190.703	.000[d]
	Residual	157629.335	479	329.080		
	Total	534168.593	485			

a. Dependent Variable: Math 2011

b. Predictors: (Constant), INC Math

c. Predictors: (Constant), INC Math, Gender, Absent, Economically Disadvantaged, Race

d. Predictors: (Constant), INC Math, Gender, Absent, Economically Disadvantaged, Race, MAth 2010

This table confirms the previous results as well as indicating that each model in and of itself is significant. The independent variables entered in Models 1, 2, and 3 predicted scores on the NJ ASK 2011 Math test to a statistically significant degree ($p<.001$) in school A (Model 1: $F=419.563$, $df= 1, 484$, $p<.001$; Model 2: $F= 92.167$, $df= 5, 480$, $p<.001$; Model 3: $F= 190.703$, $df= 6, 479$, $p<.001$).

Table 33

Hierarchical Regression Coefficients Table for School A, Mathematics

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	226.729	1.340		169.140	.000					
	INC Math	-48.309	2.358	-.681	-20.483	.000	-.681	-.681	-.681	1.000	1.000
2	(Constant)	239.625	3.074		77.959	.000					
	INC Math	-46.798	2.341	-.660	-19.990	.000	-.681	-.674	-.652	.975	1.026
	Economically Disadvantaged	-1.651	2.369	-.024	-.697	.486	-.128	-.032	-.023	.923	1.083
	Gender	-6.346	2.173	-.096	-2.920	.004	-.132	-.132	-.095	.991	1.009
	Race	-8.741	2.899	-.102	-3.015	.003	-.181	-.136	-.098	.921	1.085
	Absent	-.408	.216	-.062	-1.888	.060	-.094	-.086	-.062	.994	1.006
3	(Constant)	101.721	7.742		13.138	.000					
	INC Math	-22.854	2.195	-.322	-10.411	.000	-.681	-.430	-.258	.643	1.556
	Economically Disadvantaged	-.596	1.804	-.009	-.330	.741	-.128	-.015	-.008	.922	1.084
	Gender	-3.590	1.661	-.054	-2.161	.031	-.132	-.098	-.054	.983	1.017
	Race	-.743	2.248	-.009	-.330	.741	-.181	-.015	-.008	.888	1.126
	Absent	-.147	.165	-.022	-.894	.372	-.094	-.041	-.022	.987	1.013
	MAth 2010	.566	.030	.594	18.685	.000	.797	.649	.464	.609	1.641

a. Dependent Variable. Math 2011

The coefficients table provides a detailed analysis of the strength of each individual predictor. In model 1, the predictor of INC Math is statistically significant, $p<.001$ with t= -20.483 and β= -.681. Independently, the Beta -.681 for INC Math is a moderate-strong predictor of student performance on the 2011 NJASK assessment in mathematics, contributing 46% to the overall variance in Math 2011 performance for this model. Because the Beta is negative, this indicates that non-disabled students that are not in inclusion are doing better than non-disabled students assigned to an inclusion classroom.

When the variables of economically disadvantaged, gender, race, and student absences were added to INC MATH, the strength of the variable INC MATH decreased by only .02, or 2% (-.681 vs. -.660), meaning that these variables have a minimal effect on the strength of INC

MATH, as β= -.660 is still a moderate predictor of student performance on the NJ ASK 2011 in Mathematics and contributes between 44%-45% of the overall variance in Math 2011 for this model. It is important to note that although INC MATH was the strongest predictor in model 2, other predictors were very weak predictors; however, they were found to be statistically significant.

Gender was found to be a statistically significant variable (β=-.096, t=-2.920, *p*<.004). With a Beta -.096, gender is a weak predictor of student performance on the NJ ASK in mathematics, contributing between .9% to 1.7% of the overall variance in Math 2011 for this model. Also, race was found to be a statistically significant predictor of student performance on the NJ ASK in mathematics (β=-.102, t=-3.015, *p*<.003). The Beta for the predictor race is -.102, making it a very weak predictor of student performance on the NJ ASK in math and contributing 1.0% to 1.7% of the overall variance in Math 2011 for this model. Student absences and economically disadvantaged were not found to be statistically significant predictors of student performance on the NJ ASK in math (*p*=.060 & *p*=.486).

In Model 3, when Math 2010 (previous year's scores) were added to the existing variables, a lot of changes occurred. First, race was no longer a statistically significant variable (*p*=.741). Math 2010 was found to be a moderate statistically significant predictor (β=.594, t=18.685, *p*<.001), contributing 35% to 42% of the overall variance in Math 2011 performance for this model, meaning that non-disabled students who did well on previous year's NJ ASK, also did well on the current year's test in mathematics. Although the Beta of INC Math dropped from -.660 to -.322, .338 or 33.8%, it is still a statistically significant variable; however, its strength has decreased, becoming a weaker predictor of student performance on the NJ ASK in

math (β= -.322, t=-10.411, p<.001). In Model 3, INC Math contributes between 10.3% to 18.4% of the overall variance in Math 2011 performance in this model. The Beta for INC Math indicates that it accounts for only 10% of the variance in NJ ASK Math 2011 scores; however, the partial correlation tells us that INC Math accounts for between 10.3% to 18.4% of the variance in NJ ASK Math 2011, which indicates that the Math 2010 test scores have an influence over INC Math and are somewhat suppressing the strength of the INC Math variable.

Finally, gender is not statistically significant, but rather a notable variable, and is a weak predictor of student performance on the NJASK in math (β=.-054, t= -2.161, p<.031), contributing between .2% to.9% of the overall variance in Math 2011 performance in this model. It can be said that both the simultaneous regression as well as the hierarchical regression make one fact clear: the previous year's test scores (MATH 2010) and placement in an inclusion setting (INC MATH) are the strongest predictors of student performance on the NJ ASK in math and explain the largest proportion of the variance for School A in mathematics.

Research Questions and Answers for School A

Research Question 1: What is the influence of placement in the inclusive setting on the performance of non-disabled students in the area of mathematics as measured by the NJ ASK when controlling for student mutable variables in Grades 6, 7, and 8?

Null Hypothesis H_o^1: Placement in an inclusive setting has no significant influence on the performance of non-disabled students in Grades 6-8 in the area of mathematics as measured by the NJ ASK.

Answer: Results of this study indicate that placement in an inclusion classroom for mathematics is a statistically significant predictor of non-disabled student performance on the NJ

ASK in mathematics ($p<.001$) accounting for between 10.3% to 18.4% of the overall variance; therefore, the null hypothesis is rejected because placement in an inclusion classroom for mathematics has a statistically significant influence on student performance in the area of mathematics.

Research Question 2: What is the influence of placement in the inclusive setting on the performance for non-disabled students in Grades 6, 7, and 8 in the area of language arts and literacy when controlling for student mutable variables as measured by the NJ ASK?

Null Hypothesis H_o^2: Placement in an inclusive setting has no significant influence on the performance of non-disabled students in Grades 6-8 in the area of language arts and literacy as measured by the NJ ASK.

Answer: Results of this study indicate that placement in an inclusion classroom for language arts is a statistically significant predictor of non-disabled student performance on the NJ ASK in language arts ($p<.001$), accounting for between 10% to 18% of the overall variance. With that said, the null hypothesis is rejected because placement in an inclusion classroom for language arts has a statistically significant influence on student performance on the NJ ASK in the area of language arts.

Research Question 3: How well does placement in an inclusion classroom, student gender, attendance in class, race, and SES predict and/or influence student overall academic performance as measured by the NJ ASK 6, 7, and 8?

Null Hypothesis H_o^3: Placement in an inclusion classroom, student gender, attendance in class, race, and SES has no significant influence on the performance of non-disabled students in the areas of language arts and mathematics as measured by the NJ ASK.

Answer: Analysis of the data indicates that in the area of language arts the predictor of placement in an inclusion class for language arts (INC LA) has a negative and significant influence on general education students' performance on the NJ ASK assessment in the area of language arts (β= -.331, t=-10.390, p<.001), accounting for between 10% to 18% of the overall variance.

The predictor of LAL 2010 had a positive and significant influence on general education students' performance on the NJ ASK assessment in the area of language arts (β=.555, t=17.163, p<.001), accounting for between 31% to 38% of the overall variance.

The predictor of gender had a positive and significant influence on student performance on the NJ ASK in language arts (β=.057, t=1.956, p<.051), accounting for .3% to .7% of the overall variance. With all of this data presented, the null hypothesis is rejected because the variables of inclusion, gender, and previous achievement all have a statistically significant influence on student achievement in the area of language arts.

In the area of mathematics, the predictor of placement in an inclusion class for mathematics (INC MATH) had a negative and significant influence on general education students' performance on the NJ ASK assessment in the area of mathematics (β= -.323, t=-10.466, p=.001), accounting for between 10.3% to18.4% of the overall variance.

In the area of mathematics, the predictor of gender had a positive and significant influence on student performance on the NJ ASK 2011 in mathematics (β=-.053, t=-2.129, p<. 034), accounting for between .2% to .9% of the overall variance.

The predictor of Math 2010 had a positive and significant influence on general education students' performance on the NJ ASK assessment in the area of mathematics (β=.595, t=18.717,

p=.001), accounting for between 35% to 42% of the overall variance. Because of the evidence presented, the null hypothesis is rejected because the predictor variables of previous achievement, gender, and inclusion all have a statistically significant influence on student performance in the area of mathematics.

Language Arts, School B

Multiple regression analysis was performed to determine the amount of variability in the dependent or outcome variable, NJ ASK LAL 2011, that could be explained by the student demographic variables of (a) placement in an inclusion classroom, (b) free lunch (free lunch eligible), (c) gender, (d) race, (e) and number of days absent (absent). Table 34 provides the descriptive statistics for both the outcome variable and all predictor variables used in the model.

Table 34

Descriptive Statistics for School B, Language Arts

Descriptive Statistics

	Mean	Std. Deviation	N
LAL 2011	199.13	21.945	535
Free Lunch Eligible	.32	.465	535
Gender	.54	.499	535
Race	.63	.483	535
INCLAL	.17	.376	535
ABSENT	4.56	4.465	535
LAL2010	199.24	23.406	535

First, a correlation table (Appendix B) was examined in order to identify relationships between a dependent variable and one or more independent variables. Correlation and regression analysis are related in the sense that both deal with relationships among variables. The correlation coefficient is a measure of linear association between two variables. Values of the correlation coefficient are always between -1 and +1. A correlation coefficient of +1 indicates that two variables are perfectly related in a positive linear sense, a correlation coefficient of -1 indicates that two variables are perfectly related in a negative linear sense, and a correlation coefficient of 0 indicates that there is no linear relationship between the two variables.

With that said, the Pearson Correlation table reveals that there is a moderate negative relationship between the predictor variable of INC LA and the dependent variable of LAL 2011 (Pearson r= -.270) that was found to be statistically significant *(p<.001)*. Also, there is a strong relationship between the predictor variable LAL 2010 and the dependent variable LAL 2011 (Pearson r= .721), which was found to be statistically significant *(p<.001)*. There is expected to be a high correlation for these two variables considering that they are relatively the same exam given in two different school years, 2010 and 2011. Finally, there is a negative moderate relationship between the predictor variable of LAL 2010 and predictor variable INC LA (Pearson r=-.338), which was found to be statistically significant *(p<.001)*.

These findings indicate a need to report the partial correlation coefficients for each of these variables because one variable could on occasion act as a suppressor to the other variable when a simultaneous regression with four or more predictors is run. The partial correlation coefficient calculation partitions out the influence of highly related predictor variables on one another. In other words, there is a closer approximation of the amount of variance explained by

that particular predictor variable on the outcome variable (Leech, Barrett, & Morgan, 2008). For each analysis, the Beta is reported as well as the partial correlation coefficient, as it displays the range in percentage of overall variance accounted for in each model on the dependent variable of NJ ASK LAL 2011 test scores. Using the linear regression analysis function on SPSS, a statistically significant model emerged (See Tables 35 through 37).

Table 35

Variables Entered/Removed

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	LAL2010, ABSENT, Free Lunch Eligible, Gender, Race, INCLAL[b]	.	Enter

a. Dependent Variable: LAL 2011

b. All requested variables entered.

Table 36

Model Summary Language Arts, School B

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.730[a]	.533	.527	15.086	.533	100.325	6	528	.000

a. Predictors: (Constant), LAL2010, ABSENT, Free Lunch Eligible, Gender, Race, INCLAL

The adjusted R^2 for this model indicates that 52.7% of the variance in student performance on the NJ ASK 6-8 in language arts can be explained by student absence, gender,

placement in an inclusion classroom for language arts, race, free lunch eligibility, and the 2010 NJ ASK scores in Language Arts and Literacy. Additionally, this regression model (Table 37) is statistically significant, with F=100.325, df=6,528, p <.001.

Table 37

ANOVA for School B, Language Arts

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	137002.974	6	22833.829	100.325	.000[b]
	Residual	120172.383	528	227.599		
	Total	257175.357	534			

a. Dependent Variable: LAL 2011

b. Predictors: (Constant), LAL2010, ABSENT, Free Lunch Eligible, Gender, Race, INCLAL

Examination of the standardized coefficient (Table 38) indicates that there are two statistically significant predictors, race and LAL 2010. Multicollinearity was not an issue because the tolerance values for each variable in the model was greater than .399 (< $1 - R^2$) and all reported VIFs were close to 1.

Table 38

Coefficients for School B, Language Arts

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	73.002	6.399		11.408	.000					
	Free Lunch Eligible	-1.986	1.472	-.042	-1.349	.178	-.174	-.059	-.040	.909	1.100
	Gender	1.142	1.331	.026	.858	.391	.139	.037	.026	.967	1.034
	Race	-4.014	1.435	-.088	-2.797	.005	-.237	-.121	-.083	.886	1.128
	INCLAL	-1.309	1.857	-.022	-.705	.481	-.270	-.031	-.021	.874	1.145
	ABSENT	.177	.147	.036	1.207	.228	.015	.052	.036	.993	1.007
	LAL2010	.643	.031	.686	20.995	.000	.721	.675	.625	.830	1.205

a. Dependent Variable: LAL 2011

The predictor of race has a negative and significant influence on general education students' performance on the NJ ASK 2011 assessment in the area of language arts (β=-.088, t=-2.797, p<.005). As per the dummy coding criteria, the negative Beta -.088 indicates that general education students who are White outperform general education students who are non-White in language arts as measured by the NJ ASK. It is also important to note that the closer the Beta to 1, the stronger the influence of the predictor. The Beta -.088 indicates that although it is a predictor of student performance on the NJ ASK in language arts, it is a weak predictor contributing between .7% to 1.4% of the overall variance in LAL 2011 for this model.

The predictor of LAL 2010 had a positive and significant influence on general education students' performance on the NJ ASK assessment in the area of language arts (β=.686, t=20.995, p<.001). The higher the previous year's 2010 NJ ASK score in the area of language arts, the better they did on the NJ ASK test during the 2011 school year. Again, the closer the Beta is to 1, the stronger the influence of the predictor. The Beta of .686 indicates that it is a moderate to strong predictor of student performance on the NJ ASK 2011 in language arts, contributing between 47% to 45% to the overall variance in LAL 2011 performance for this model. Additionally, LAL 2010 has a 7.8 times greater influence on student performance on language arts than race as calculated by dividing the Beta for race .088 into the Beta for LAL 2010 .686.

The independent variables of free lunch eligible, gender, placement in an inclusion classroom for language arts, and absences were not found to be statistically significant predictors of performance on the 2011 NJ ASK LAL assessment, as the *p* values for these variables were all greater than .05.

Given the results produced by the simultaneous linear regression models, additional clarity was required in order to more thoroughly examine the r squared change between the variables. Hierarchical multiple regression is a variant of the basic multiple regression procedure that allows the researcher to specify a fixed order of entry for variables in order to control for the effects of covariates or to test the effects of certain predictors independent of the influence of others (Leech, Barrett, & Morgan, 2011). Because the outputs indicated that the previous year's test scores were the strongest predictors of student achievement on the NJ ASK in language arts, the hierarchical regression analysis allowed the variables to be entered in various models as shown below.

Table 39

Hierarchical Regression Block Inputs, School B Language Arts

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	INCLAL[b]	.	Enter
2	Gender, ABSENT, Economically Disadvantaged, Race[b]	.	Enter
3	LAL2010[b]	.	Enter

a. Dependent Variable: LAL 2011

b. All requested variables entered.

As displayed in Table 39, variables were entered into the regression models as per the following blocks: Model 1, Inclusion LA; Model 2, Inclusion LA, Absent, Gender, Economically Disadvantaged, Race; Model 3, Inclusion LA, Absent, Gender, Economically Disadvantaged, Race, LAL 2010 (previous year's scores). By entering the variables in block models, any observed effect on LAL 2011 scores can then be said to be "independent of" the effects of the variables previously controlled for.

Table 40

Model Summary Hierarchical Regression, School B, Language Arts

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.270[a]	.073	.071	21.151	.073	41.885	1	533	.000
2	.372[b]	.139	.130	20.465	.066	10.079	4	529	.000
3	.729[c]	.532	.526	15.101	.393	443.527	1	528	.000

a. Predictors: (Constant), INCLAL

b. Predictors: (Constant), INCLAL, Gender, ABSENT, Economically Disadvantaged, Race

c. Predictors: (Constant), INCLAL, Gender, ABSENT, Economically Disadvantaged, Race, LAL2010

It is also essential to examine the R squared change, which examines the percentage of variability in the dependent variable that can be accounted for by all the predictors together. The change in R squared is a way to evaluate how much predictive power was added to the model by the addition of another variable. The hierarchical regression method will also determine if the R squared change is significant as each block model is introduced. In examination of the model summary, for Model 1, the R squared change began at .073. The analysis of the adjusted R squared indicates INC LA accounts for approximately 7.1% of the variability, which is deemed statistically significant ($p<.001$). When the additional variables are entered into Model 2, the R

squared change is minimal, .066, and the percentage of variability (adjusted R squared) accounted for changes from .071 (7.1%) to .130 (13.0%) or 5.9%. However, when Model 3 is added, which adds the previous year's test scores, there is an R squared change of .393, and the adjusted R squared is .526, meaning that 52.6% of the variance is now accounted for when all of the variables are entered into the regression.

Table 41

Hierarchical Regression ANOVA Table for School B, Language Arts

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	18737.289	1	18737.289	41.885	.000[b]
	Residual	238438.068	533	447.351		
	Total	257175.357	534			
2	Regression	35622.383	5	7124.477	17.011	.000[c]
	Residual	221552.974	529	418.815		
	Total	257175.357	534			
3	Regression	136767.026	6	22794.504	99.956	.000[d]
	Residual	120408.331	528	228.046		
	Total	257175.357	534			

a. Dependent Variable: LAL 2011

b. Predictors: (Constant), INCLAL

c. Predictors: (Constant), INCLAL, Gender, ABSENT, Economically Disadvantaged, Race

d. Predictors: (Constant), INCLAL, Gender, ABSENT, Economically Disadvantaged, Race, LAL2010

This table confirms the previous results. The independent variables entered in Models 1, 2, and 3 predicted scores on the NJ ASK LAL 2011 test in language arts to a statistically

significant degree (p<.001) in school B (Model 1: F=41.885, df= 1, 533, p<.001 ; Model 2: F= 17.011, df= 5, 529, p<.001; Model 3: F= 99.956, df= 6, 528, p<.001).

Table 42

Hierarchical Regression Coefficients Table for School B, Language Arts

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	201.806	1.004		201.049	.000					
	INCLAL	-15.751	2.434	-.270	-6.472	.000	-.270	-.270	-.270	1.000	1.000
2	(Constant)	203.942	2.012		101.387	.000					
	INCLAL	-13.775	2.388	-.236	-5.768	.000	-.270	-.243	-.233	.972	1.028
	Economically Disadvantaged	-3.828	1.894	-.087	-2.021	.044	-.162	-.088	-.082	.885	1.129
	Gender	6.105	1.777	.139	3.435	.001	.139	.148	.139	.998	1.002
	Race	-7.837	1.970	-.173	-3.978	.000	-.237	-.170	-.161	.865	1.156
	ABSENT	.188	.199	.038	.944	.346	.015	.041	.038	.995	1.005
3	(Constant)	72.638	6.409		11.334	.000					
	INCLAL	-1.300	1.859	-.022	-.699	.485	-.270	-.030	-.021	.874	1.145
	Economically Disadvantaged	-1.241	1.403	-.028	-.884	.377	-.162	-.038	-.026	.879	1.138
	Gender	1.185	1.332	.027	.889	.374	.139	.039	.026	.967	1.034
	Race	-4.097	1.465	-.090	-2.797	.005	-.237	-.121	-.083	.852	1.173
	ABSENT	.167	.147	.034	1.135	.257	.015	.049	.034	.995	1.005
	LAL2010	.645	.031	.688	21.060	.000	.721	.676	.627	.832	1.202

a. Dependent Variable. LAL 2011

The coefficients table provides a detailed analysis of the strength of each individual predictor. In Model 1, the predictor of INC LA is statistically significant, p<.001 with t= -6.472 and a β= -.270. Independently, INC LA is a weak predictor of student performance on the NJ ASK assessment in language arts, contributing between 7.2% to the overall variance in LAL 2011 performance for this model. Additionally, because the Beta is negative, this indicates that non-disabled students that are not in inclusion are doing better than non-disabled students assigned to an inclusion classroom.

When the variables of economically disadvantaged, gender, race, and student absences were added to INC LA, the strength of the variable INC LA decreases by only .04. (-.270 vs. -.236), meaning that these variables have a minimal effect on the strength of INC LA, as β= -.236

is still a weak predictor of student performance, contributing between 5.5% to 5.9% of the overall variance in LAL 2011 performance for this model.

It is important to note that although INC LA was the strongest predictor in Model 2, other predictors were very weak predictors; however, they were found to be statistically significant. Economically disadvantaged was found to be a statistically significant variable (β= -.087, t= -5.768, p<.044). With a Beta -.087, economically disadvantaged is a weak predictor of student performance on the NJ ASK in language arts contributing between .7% to .77% of the overall variance in LAL 2011 performance for this model.

Gender was also found to be a statistically significant variable (β=.139, t=3.435, p<.001). With a Beta .139, gender is a weak predictor of student performance on the NJ ASK in language arts, contributing between 1.9% to 2.2% of the overall variance in LAL 2011 performance for this model.

Also, race was found to be a statistically significant predictor of student performance on the NJ ASK in language arts (β=-.173, t=-3.978, p<.001). The Beta for the predictor race is -173, making it a very weak predictor of student performance on the NJ ASK in language arts, contributing between 2.89% to 2.9% to the overall variance in LAL 2011 performance for this model.

Student absence was not found to be a statistically significant predictor of student performance on the NJ ASK in language arts (p>.05).

In Model 3, when LA 2010 (previous year's scores) were added to the existing variables, changes occurred. First, economically disadvantaged, gender, and INC LA were no longer statistically significant variables (p=.377, p=.485 & p=.374). LA 2010 was found to be a

moderate-strong statistically significant predictor (β=.688, t=21.060, $p<.001$), contributing between 45% to 47% of the overall variance in LAL 2011 for this model, meaning that non-disabled students who did well on the previous year's NJASK, also did well on the current year's test in language arts.

Finally, race is a statistically significant variable and is a weak predictor of student performance on the NJ ASK in language arts. (β=-.090, t= -2.797, $p<.005$), contributing between .8% to 1.4% of the overall variance in LAL 2011 for this model. It can be said that both the simultaneous regression as well as the hierarchical regression make one fact clear: the previous year's test scores (LA 2010) and race are the strongest predictors of student performance on the NJ ASK in language arts and explain the largest proportion of the variance for School B in language arts.

Mathematics, School B

Multiple regression analysis was performed to determine the amount of variability in the dependent or outcome variable, N JASK Math 2011, that could be explained by the student demographic variables of (a) placement in an inclusive classroom, (b) free lunch (free lunch eligible), (c) gender, (d) race, and (e) number of days absent (absent). Table 43 provides the descriptive statistics for both the outcome variable and all predictor variables used in the model.

Table 43

Descriptive Statistics Mathematics, School B

Descriptive Statistics

	Mean	Std. Deviation	N
Math 2011	216.01	33.982	535
Gender	.54	.499	535
Race	.63	.483	535
ABSENT	4.56	4.465	535
INCMath	.14	.351	535
Math 2010	219.60	34.628	535
Free Lunch Eligible	.32	.465	535

First, a correlation table (Appendix B) was examined in order to identify relationships between a dependent variable and one or more independent variables. Correlation and regression analysis are related in the sense that both deal with relationships among variables.

With that said, the Pearson Correlation table reveals that there is a strong negative relationship between the predictor variable of INC Math and the dependent variable of Math 2011 (Pearson r= -.355) that was found to be statistically significant (*p<.001)*. Also, there is a strong relationship between the predictor variable Math 2010 and the dependent variable Math 2011 (Pearson r= .826), which was found to be statistically significant (*p<.001)*. There is expected to be a high correlation for these two variables considering that they are relatively the same exam given in two different school years 2010 & 2011. Last, there is a negative moderate relationship between the predictor variable of Math 2010 and predictor variable INC Math (Pearson r=-.411), which was found to be statistically significant (*p<.001)*

These findings indicate a need to report the partial correlation coefficients for each of these variables because one variable could on occasion act as a suppressor to the other variable when a simultaneous regression with 4 or more predictors is run. The partial correlation coefficient calculation partitions out the influence of highly related predictor variables on one another. In other words, there is a closer approximation of the amount of variance explained by

that particular predictor variable on the outcome variable (Leech, Barrett, & Morgan, 2008). For each analysis, the Beta is reported as well as the partial correlation coefficient, as it displays the range in percentage of overall variance accounted for in each model on the dependent variable of NJ ASK Math 2011 test scores. Using the linear regression analysis function on SPSS, a statistically significant model emerged (See Tables 44 through 46).

Table 44

Entered/Removed Variables

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Free Lunch Eligible, Gender, ABSENT, INCMath, Race, Math 2010[b]	.	Enter

a. Dependent Variable: Math 2011

b. All requested variables entered.

Table 45

Model Summary Math, School B

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.828[a]	.685	.681	19.181	.685	191.342	6	528	.000

a. Predictors: (Constant), Free Lunch Eligible, Gender, ABSENT, INCMath, Race, Math 2010

The adjusted R^2 for this model indicates that 68.1% of the variance in student performance on the NJ ASK 6-8 in math can be explained by student absence, gender, and placement in an inclusion classroom for language arts, race, free lunch eligibility, and the 2010

NJ ASK scores in math. Additionally, this regression model (Table 46) is statistically significant, with $F=191.342$, $df=6,528$, $p <.001$.

Table 46

ANOVA Math, School B

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	422380.599	6	70396.767	191.342	.000[b]
	Residual	194256.384	528	367.910		
	Total	616636.983	534			

a. Dependent Variable: Math 2011

b. Predictors: (Constant), Free Lunch Eligible, Gender, ABSENT, INCMath, Race, Math 2010

Examination of the standardized coefficient (Table 47) indicates that there is only one statistically significant predictor, Math 2010. Multicollinearity was not an issue because the tolerance values for each variable in the model was greater than .399 ($< 1 - R^2$) and all reported VIFs were close to 1, for the predictors were not exceedingly low ($<1-R^2$).

Table 47

Standardized Coefficient Table Mathematics, School B

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	42.056	6.869		6.122	.000					
	Free Lunch Eligible	.582	1.878	.008	.310	.757	-.173	.013	.008	.902	1.108
	Math 2010	.796	.028	.811	28.946	.000	.826	.783	.707	.760	1.317
	INCMath	-1.601	2.599	-.017	-.616	.538	-.355	-.027	-.015	.826	1.210
	ABSENT	.027	.187	.004	.142	.887	-.071	.006	.003	.986	1.014
	Race	-2.719	1.854	-.039	-1.467	.143	-.282	-.064	-.036	.858	1.165
	Gender	1.396	1.671	.020	.835	.404	-.032	.036	.020	.991	1.009

a. Dependent Variable. Math 2011

In this model there was only one statistically significant predictor. The predictor of Math 2010 had a positive and significant influence on general education students' performance on the NJ ASK assessment in the area of mathematics ($\beta=.811$, $t=28.946$, $p<.001$). The higher the previous year's 2010 NJ ASK score in the area of mathematics, the better they did on the NJ ASK test during the 2011 school year. Again, the closer the Beta is to 1, the stronger the influence of the predictor. The Beta .811 indicates that it is a strong predictor of student performance on the NJ ASK in mathematics, contributing between 61.3% to 66% of the overall variance in Math 2011 performance for this model. In this particular model, previous achievement outweighed all other variables.

The independent variables of placement in inclusion for mathematics (INC Math), free lunch eligible, gender, race, and absences were not found to be statistically significant predictors of performance on the 2011 NJ ASK Math assessment as the *p* values for these variables were all greater than .05.

Given the results produced by the simultaneous linear regression models, additional clarity was required in order to more thoroughly examine the R squared change between the variables. Hierarchical multiple regression is a variant of the basic multiple regression procedure that allows the researcher to specify a fixed order of entry for variables in order to control for the

effects of covariates or to test the effects of certain predictors independent of the influence of others (Leech, Barrett, & Morgan, 2011). Because the outputs indicated that the previous years test scores were the strongest predictors of student achievement on the NJ ASK 2011 in mathematics, the hierarchical regression analysis allowed the variables to be entered in various models as shown below.

Table 48

Hierarchical Regression Variables Entered, Mathematics School B

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	INCMath[b]	.	Enter
2	Gender, Economically Disadvantaged, ABSENT, Race[b]	.	Enter
3	Math 2010[b]	.	Enter

a. Dependent Variable: Math 2011

b. All requested variables entered.

As displayed in Table 48, variables were entered into the regression models as per the following blocks: Model 1, Inclusion Math; Model 2, Inclusion Math, Absent, Gender, Economically Disadvantaged, Race; Model 3, Inclusion Math, Absent, Gender, Economically Disadvantaged, Race, Math 2010 (previous year's scores). By entering the variables in block

models, any observed effect on Math 2011 scores can then be said to be "independent of" the effects of the variables previously controlled for.

Table 49

Model Summary Hierarchical Regression, School B Mathematics

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.355[a]	.126	.125	31.794	.126	77.022	1	533	.000
2	.426[b]	.181	.174	30.890	.055	8.913	4	529	.000
3	.828[c]	.685	.681	19.181	.504	843.966	1	528	.000

a. Predictors: (Constant), INCMath

b. Predictors: (Constant), INCMath, Gender, Economically Disadvantaged, ABSENT, Race

c. Predictors: (Constant), INCMath, Gender, Economically Disadvantaged, ABSENT, Race, Math 2010

It is also essential to examine the R squared change which examines the percentage of variability in the dependent variable that can be accounted for by all the predictors together. The change in R squared is a way to evaluate how much predictive power was added to the model by the addition of another variable. The hierarchical regression method will also determine if the R squared change is significant as each block is introduced. In examination of the model summary, for Model 1, the R squared change began at .126. The analysis of the adjusted R squared indicates, INC Math accounts for approximately 12.5% of the variability, which is deemed statistically significant ($p<.001$). When the additional variables are entered in Model 2, the R squared change is minimal, .055, and the percentage of variability (adjusted R squared) accounted for changes from .125 (12.5%) to .174 (17.4%), or 4.9%. Although a significant change, this translates into Model 2 explaining only a very small proportion of the variance. However, when Model 3 is added, which adds the previous year's test scores, there is an R

squared change of .504, and the adjusted R squared is .681, meaning that 68.1% of the variance is now accounted for when all of the variables are entered into the regression.

Table 50

Hierarchical Regression ANOVA Table for School B, Mathematics

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	77857.339	1	77857.339	77.022	.000[b]
	Residual	538779.645	533	1010.844		
	Total	616636.983	534			
2	Regression	111877.366	5	22375.473	23.450	.000[c]
	Residual	504759.617	529	954.177		
	Total	616636.983	534			
3	Regression	422380.633	6	70396.772	191.342	.000[d]
	Residual	194256.350	528	367.910		
	Total	616636.983	534			

a. Dependent Variable: Math 2011

b. Predictors: (Constant), INCMath

c. Predictors: (Constant), INCMath, Gender, Economically Disadvantaged, ABSENT, Race

d. Predictors: (Constant), INCMath, Gender, Economically Disadvantaged, ABSENT, Race, Math 2010

This ANOVA table confirms the previous results. The independent variables entered in Models 1, 2, and 3 predicted scores on the NJ ASK 2011 test in mathematics to a statistically significant degree ($p<.001$) in school B (Model 1: $F=77.022$, df= 1, 533, $p<.001$; Model 2: $F= 23.450$, df 5, 529, $p<.001$; Model 3: $F= 191.342$, df= 6, 479, $p<.001$).

Table 51

Hierarchical Regression Coefficients Table for School B, Mathematics

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Correlations Zero-order	Correlations Partial	Correlations Part	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	220.952	1.486		148.727	.000					
	INCMath	-34.368	3.916	-.355	-8.776	.000	-.355	-.355	-.355	1.000	1.000
2	(Constant)	233.528	2.976		78.463	.000					
	INCMath	-29.887	3.869	-.309	-7.724	.000	-.355	-.318	-.303	.962	1.039
	Free Lunch Eligible	-6.138	2.995	-.084	-2.050	.041	-.173	-.089	-.080	.916	1.091
	Gender	-2.052	2.678	-.030	-.766	.444	-.032	-.033	-.030	.997	1.004
	Race	-14.335	2.908	-.204	-4.930	.000	-.282	-.210	-.193	.901	1.110
	ABSENT	-.251	.300	-.033	-.835	.404	-.071	-.036	-.033	.989	1.011
3	(Constant)	42.056	6.869		6.122	.000					
	INCMath	-1.601	2.599	-.017	-.616	.538	-.355	-.027	-.015	.826	1.210
	Free Lunch Eligible	.582	1.878	.008	.310	.757	-.173	.013	.008	.902	1.108
	Gender	1.396	1.671	.020	.835	.404	-.032	.036	.020	.991	1.009
	Race	-2.719	1.854	-.039	-1.467	.143	-.282	-.064	-.036	.858	1.165
	ABSENT	.027	.187	.004	.142	.887	-.071	.006	.003	.986	1.014
	Math 2010	.796	.028	.811	28.946	.000	.826	.783	.707	.760	1.317

a. Dependent Variable: Math 2011

The coefficients table provides a detailed analysis of the strength of each individual predictor. In Model 1, the predictor of INC MATH is statistically significant, $p<.001$ with t= -8.776 and a β= -.355. Independently, INC MATH is a weak predictor of student performance on the NJ ASK assessment in mathematics, accounting for 12.6% of the overall variance in Math 2011 performance for this model. Additionally, because the Beta is negative, this indicates that non-disabled students that are not in inclusion are doing better than non-disabled students assigned to an inclusion classroom.

When the variables of economically disadvantaged, gender, race, and student absences were added to INC MATH, the strength of the variable INC Math decreased by only .04 or 4%. (-.355 vs. -.312), meaning that these variables have a minimal effect on the strength of INC

Math, as β= -.312 is still a weak predictor of student performance on the NJ ASK 2011 in mathematics, contributing between 9.7% to 10.1% of the overall variance in Math 2011 performance for this model. It is important to note that although INC Math was the strongest predictor in Model 2, there was one other predictor that was a very weak predictor; however, it was found to be statistically significant. Race was found to be a statistically significant predictor of student performance on the NJ ASK in mathematics (β=-.208, t=-4.907, *p*<.001). The Beta for the predictor race is -.208, making it a very weak predictor of student performance on the NJ ASK in math, contributing between 4.3% to 4.4% of the overall variance in Math 2011 performance for this model. Student absences, economically disadvantaged, and race were not found to be statistically significant predictors of student performance on the NJASK 2011 in math (*p*=.176; *p*=.485, *p*=.349).

In Model 3, when Math 2010 (previous year's scores) were added to the existing variables, changes occurred. First, race and INC math were no longer statistically significant variables (*p*=.539, *p*=.205). Math 2010 was found to be a strong statistically significant predictor, contributing between 61.3% to 66% of the overall variance in Math 2011 performance for this model (β=.810, t=29.051, *p*<.001), meaning that non-disabled students who did well on previous years NJ ASK also did well on the current year's test in math. It can be said that both the simultaneous regression as well as the hierarchical regression make one fact clear: the previous year's test scores (Math 2010) are the strongest predictors of student performance on the NJ ASK 2011 in math and explains the largest proportion of the variance for School B in math.

Research Questions and Answers for School B

Research Question 1: What is the influence of placement in the inclusive setting on the performance of non-disabled students in the area of mathematics as measured by the NJ ASK when controlling for student mutable variables in Grades 6, 7, and 8?

Null Hypothesis: H_0^1: Placement in an inclusive setting has no statistically significant influence on the performance of non-disabled students in Grades 6-8 in the area of mathematics as measured by the NJ ASK 2011.

Answer: Results of this study indicate that placement in an inclusion classroom for mathematics was not a statistically significant predictor of non-disabled student performance on the NJ ASK in mathematics ($p > .05$); therefore, the null hypothesis is not rejected because placement in an inclusion classroom for mathematics did not have a statistically significant influence on student performance in the area of mathematics.

It is important to note that in School B, the previous year's test scores (Math 2010) presented a situation where the partial correlation coefficient actually decreased the percentage of the variance accounted for. In examination of the hierarchical regression, Model 3, the Beta for Math 2010 was .811 (66% of the variance) but the partial coefficient was actually less, .783 (61% of the variance). Thus the partial coefficient is actually bringing the percentage of the variance accounted for from 66% to 61%, which means that if Math 2010 (previous year's test scores) could possibly account for 5% less of the variance, then other variables could perhaps account for more of the variance. This is one possibility as to why the other variables in School B Mathematics were not found to be statistically significant. Math 2010 was acting as a suppressor variable.

Research Question 2: What is the influence of placement in the inclusive setting on the performance for non-disabled students in Grades 6, 7, and 8 in the area of language arts and literacy when controlling for student mutable variables as measured by the NJ ASK?

Null Hypothesis H_0^2: Placement in an inclusive setting has no statistically significant influence on the performance of non-disabled students in Grades 6-8 in the area of language arts and literacy as measured by the NJ ASK.

Answer: Results of this study indicate that placement in an inclusion classroom for language arts is not a statistically significant predictor of non-disabled student performance on the NJ ASK in language arts (p>.05). With that said, the null hypothesis is not rejected because placement in an inclusion classroom for language arts does not have a statistically significant influence on student performance on the NJASK 2011 in the area of language arts.

It is important to note that in School B, the previous year's test scores (LA 2010) presented a situation where the partial correlation coefficient actually decreased the percentage of the variance accounted for. In examination of the hierarchical regression, Model 3, the Beta for LA 2010 was .688 (47.3% of the variance), but the partial coefficient was actually less, .627 (45% of the variance). Thus the partial coefficient is actually bringing the percentage of the variance accounted for from 47.3% to 45%, which means that if LA 2010 (previous year's test scores) could possibly account for approximately 2% less of the variance, then other variables could perhaps account for more of the variance. This is one possibility as to why the other variables in School B language arts were not found to be statistically significant. LA 2010 was acting as a suppressor variable.

Research Question 3: How well does placement in an inclusion classroom, student gender, attendance in class, race, and SES predict and/or influence student overall academic performance as measured by the NJ ASK 6,7, and 8?

Null Hypothesis H_o^3: Placement in an inclusion classroom, student gender, attendance in class, race, and SES have no statistically significant influence on the performance of non-disabled students in the areas of language arts and mathematics as measured by the NJ ASK 2011.

Answer: Analysis of the data indicates that in the area of language arts the predictor of race has a negative and significant influence on general education students' performance on the NJ ASK assessment in the area of language arts (β=-.088, t=-2.797, $p \leq .005$) accounting for between .8% to 1.4% of the overall variance. The predictor of LAL 2010 has a positive and significant influence on general education students' performance on the NJ ASK assessment in the area of language arts (β=.686, t=20.995, $p < .001$), accounting for between 45% to 47% of the overall variance.

In the area of mathematics, there was only one statistically significant predictor. The predictor of Math 2010 had a positive and significant influence on general education students' performance on the NJASK assessment in the area of mathematics (β=.811, t=28.946, $p = .001$) accounting for between 61.3%-66% of the overall variance. In this particular model, previous achievement out weighted all other variables.

The independent variables of placement in inclusion for mathematics (INC Math), free lunch eligible, gender, race, and absences were not statistically significant predictors of performance on the 2011 NJ ASK assessment, as the *p* value was greater than .05. Because of

the evidence presented, the null hypothesis is not rejected because the predictor variables do not have a statistically significant influence on student achievement in the area of mathematics.

It is important to note that in School B, both previous year's test scores (LA 2010 and Math 2010) presented a situation where the partial correlation coefficient actually decreased the percentage of the variance accounted for. This is one possibility as to why the other variables in School B were not found to be statistically significant. LA 2010 and Math 2010 were acting as suppressor variables and interacting with some of the other variables, possibly ones that are not even significant in these models.

Factorial ANOVA Schools A and B, Language Arts and Literacy

Given the results of the multiple and hierarchical regression models which suggest that in School A, inclusion is a statistically significant variable that influences general education student performance on the NJ ASK in Grades 6, 7, and 8 and in School B, inclusion is not a statistically significant variable that influences general education student performance, further statistical analysis was conducted in order to determine if there was a significant interaction between inclusion in the two schools.

Factorial ANOVA was chosen in order to compare groups based on two independent variables (inclusion and school code) while controlling for previous achievement scores. Additionally, it is important to note that because previous achievement scores (LA 2010, Math 2010) accounted for the largest percentage of variance in each school, they were inputted as a control in the factorial ANOVA analysis.

Table 52

Test of Between Subject Effects School A and B, Language Arts

Tests of Between-Subjects Effects

Dependent Variable:LAL 2011

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	393563.107[a]	4	98390.777	396.233	.000	.577	1584.933	1.000
Intercept	78487.034	1	78487.034	316.078	.000	.214	316.078	1.000
LAL2010	329458.943	1	329458.943	1326.777	.000	.533	1326.777	1.000
incLAL	6874.517	1	6874.517	27.685	.000	.023	27.685	1.000
SchoolCode	4.307	1	4.307	.017	.895	.000	.017	.052
incLAL * SchoolCode	4495.496	1	4495.496	18.104	.000	.015	18.104	.989
Error	288293.999	1161	248.315					
Total	45327251.000	1166						
Corrected Total	681857.105	1165						

a. R Squared = .577 (Adjusted R Squared = .576)

b. Computed using alpha = .05

Estimated Marginal Means

School Code

Estimates

Dependent Variable:LAL 2011

School Code	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
MMS	194.456[a]	.835	192.818	196.095
SMS	194.310[a]	.740	192.857	195.763

a. Covariates appearing in the model are evaluated at the following values: LAL2010 = 194.35.

In the ANOVA analysis (See Table 52), the influence of previous achievement (LAL 2010) was found to have a significant impact on the dependent variable of LAL 2011 (F=1326.777, df= 1,1161, p≤.000) The Eta (index for the effect size for each independent variable and the interaction) for LAL 2010 is .533, meaning that 53.3% in language arts achievement can be predicted by prior achievement. Also, the influence of inclusion on language arts (INC LA) was found to have a significant impact on the dependent variable of LAL 2011

($F=27.685$, $df= 1,1161$, $p \leq .000$). The Eta (index for the effect size for each independent variable and the interaction) for INC LA is .023, meaning that 2.3% in language arts achievement can be predicted by INC LA. School code was not found to have a significant influence on the dependent variable of LA 2011 ($p=.895$) It is important to note that school code, as an independent variable, was not found to have a significant impact on the dependent variable; however, it was found to have an impact in its interaction with the independent variable of inclusion (INC LA). Finally, in the ANOVA output the interaction between INC LA and school code was found to have a significant impact on the dependent variable of LA 2011 ($F=18.104$, $df= 1, 1161$, $p \leq .000$). The Eta (index for the effect size for each independent variable and the interaction) for the interaction between School Code and INC LA is .015, meaning that 1.5% in language arts achievement can be predicted by the interaction between inclusion and school code.

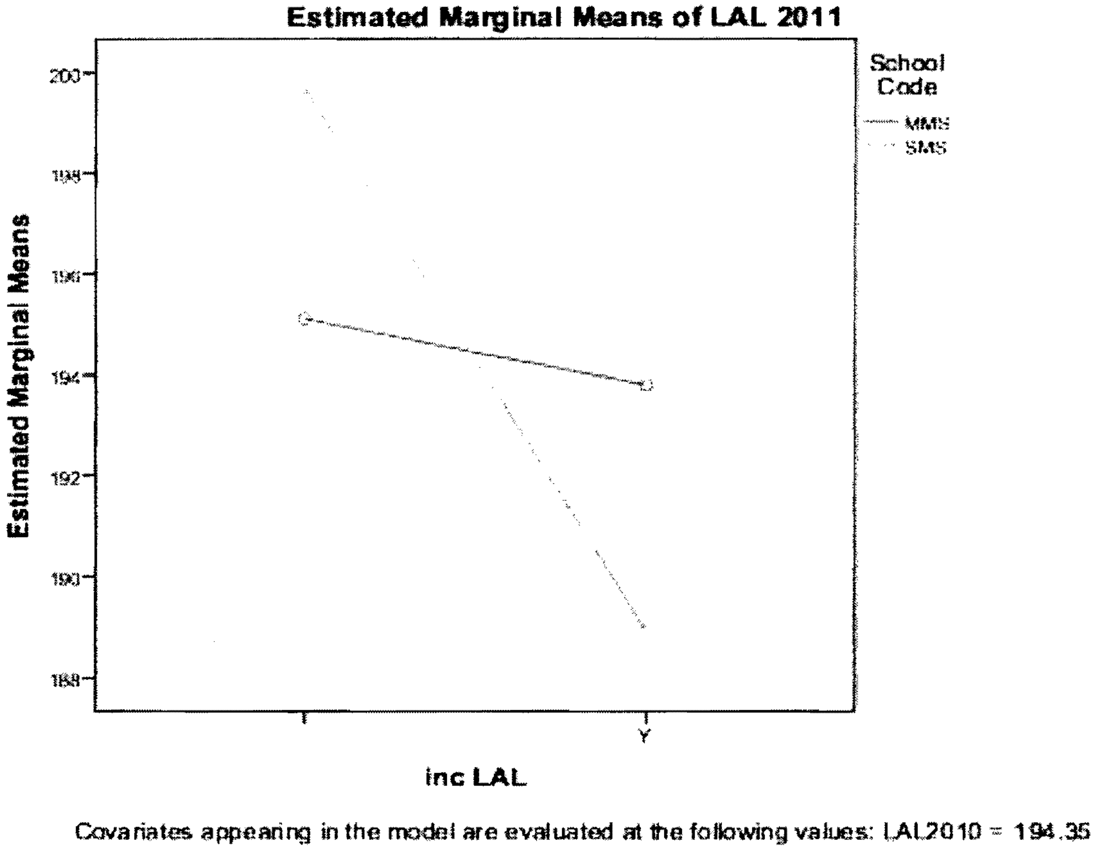


Figure 2. Estimated marginal means of LAL 2011.

Examination of the profile plot displays a disordinal interaction, which indicates that there is a statistically significant interaction. In School A (SMS) there is a large difference in the general education student test scores for students placed in inclusion and those not placed in inclusion. This analysis suggests that the general education students placed in inclusion classes in School A are more homogeneously grouped by lower NJ ASK scores in the area of language arts. In School B (MMS), there is some indication that the general education students in inclusion have lower test scores, but the difference between the students in and not in inclusion is not as

dramatic as in School A. It can be said that the effect of LA 2011 scores depends on which school a general education student is placed in an inclusion class.

In summary, the difference between LAL scores is significant in both schools based on inclusion status, when controlling for the previous year's NJ ASK (LA 2010) scores. When the interaction between schools is examined--in other words, when there is more of a significant difference in scores between inclusion and non-inclusion students based on the school the student attends--the results are drastic. There is a significant and much larger difference in LAL 2011 performance between inclusion and non-inclusion students based on school. Students in MMS, although significantly different, do not have as severe a drop across inclusion status as students in SMS. This analysis supports the idea that one school (SMS) places low performing general education students in inclusion classes in the area of language arts.

Factorial ANOVA Schools A and B, Math

Table 53

Tests of Between-Subjects Effects

Tests of Between-Subjects Effects

Dependent Variable: Math2011

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	987053.052[a]	4	246763.263	556.714	.000	.656	2226.855	1.000
Intercept	59066.667	1	59066.667	133.258	.000	.102	133.258	1.000
Math2010	752839.752	1	752839.752	1698.455	.000	.593	1698.455	1.000
SchoolCode	495.930	1	495.930	1.119	.290	.001	1.119	.185
IncMath	6790.506	1	6790.506	15.320	.000	.013	15.320	.974
SchoolCode * IncMath	4534.345	1	4534.345	10.230	.001	.009	10.230	.892
Error	517272.497	1167	443.250					
Total	52371270.000	1172						
Corrected Total	1504325.549	1171						

a. R Squared = .656 (Adjusted R Squared = .655)

b. Computed using alpha = .05

Estimated Marginal Means

School Code

Estimates

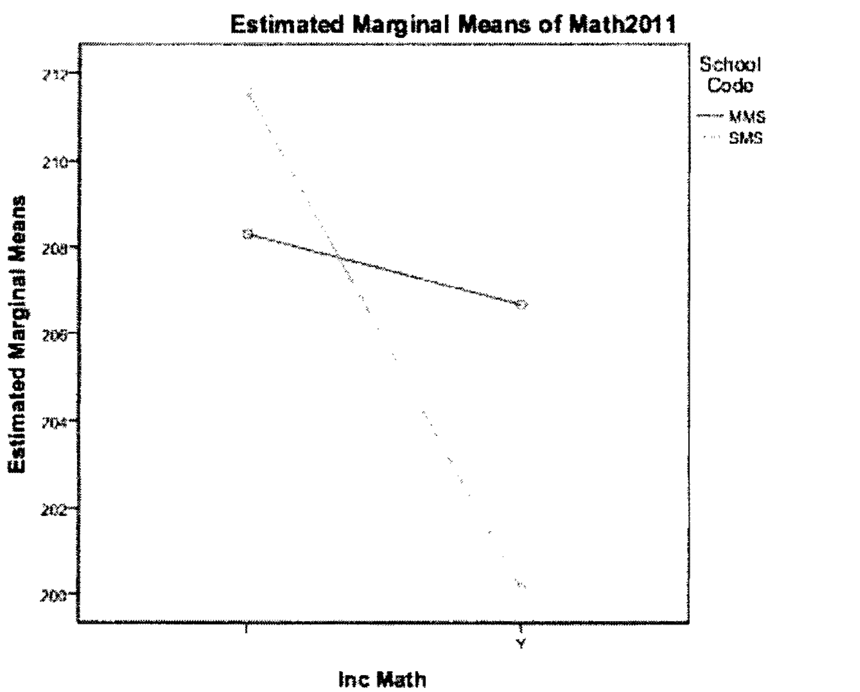
Dependent Variable: Math2011

School Code	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
MMS	207.478[a]	1.155	205.212	209.745
SMS	205.856[a]	1.046	203.804	207.908

a. Covariates appearing in the model are evaluated at the following values: Math2010 = 212.15.

In the ANOVA Analysis (See Table 52) the influence of previous achievement (Math 2010) was found to have a significant impact on the dependent variable of Math 2011 (F=1698.455, df= 1,1167, p≤.000) The Eta (index for the effect size for each independent variable and the interaction) for Math 2010 is .593, meaning that 59.3% in math achievement can be predicted by prior achievement.. Also, the influence of inclusion math (INC MATH) was found to have a significant impact on the dependent variable of Math 2011 (F=15.320, df= 1,1167, p≤.000). The Eta (index for the effect size for each independent variable and the

interaction) for INC Math is .013, meaning that 1.3% in math achievement can be predicted by INC Math. School code was not found to have a significant influence on the dependent variable of Math 2011 (p=.290). It is important to note that school code, as an independent variable, was not found to have a significant impact on the dependent variable; however, it was found to have an impact in its interaction with the independent variable of inclusion (INC Math). Finally, in the ANOVA output the interaction between INC Math and school code was found to have a significant impact on the dependent variable of Math 2011 (F=10.230, df= 1, 1167, p≤.001). The Eta (index for the effect size for each independent variable and the interaction) for the interaction between INC Math and school code is .009, meaning that 0.9% in math achievement can be predicted by INC Math and school code.



Covariates appearing in the model are evaluated at the following values: Math2010 = 212.15

Figure 3. Estimated marginal means of Math 2011.

Examination of the profile plot displays a disordinal interaction, which indicates that the interaction is statistically significant. In School A (SMS) there is a large difference in the general education student test scores for students placed in inclusion and those not placed in inclusion. This analysis suggests that the general education students placed in inclusion classes in School A are more homogeneously grouped by lower NJ ASK scores in the area of math. In School B (MMS), there is some indication that the general education students in inclusion have lower test scores, but the difference between the students in and not in inclusion is not as dramatic as in School A.

In summary, the difference between math scores is significant in both schools based on inclusion status, when controlling for the previous year's NJ ASK (Math 2010) scores. When the interaction between schools is examined--in other words, when there is more of a significant difference in scores between inclusion and non-inclusion students based on the school the student attends--the results suggest a large difference. There is a significant and much larger difference in Math 2011 performance between inclusion and non-inclusion students based on school. Students in MMS, although significantly different, do not have as severe a drop across inclusion status as students in SMS. This analysis supports the idea that one school (SMS) places low performing general education students in inclusion classes in the area of math.

Chapter V

CONCLUSIONS AND RECOMMENDATIONS

Introduction

The current public education system in New Jersey continues to be driven by two large pieces of legislation, No Child Left Behind (NCLB) and Individual with Disabilities Improvement Act (IDEIA). NCLB emphasizes accountability for administrators and teachers, as it requires 100% proficiency on standardized testing for all students by 2014, deeming each child in New Jersey proficient in the content areas of language arts and mathematics. In addition, IDEIA mandates that all children with disabilities receive a free and appropriate education in the least restrictive environment and that the movement toward inclusion of students with disabilities in the general education classroom dominates and progresses the field of education. As the demand for proficiency increases annually and inclusion becomes standard practice, the lack of empirical quantitative evidence showing the influence of the inclusive classroom on the academic achievement of non-disabled students becomes more apparent.

The purpose for this study was to determine whether placement in an inclusive setting affects the academic achievement of general education students on the Language Arts Literacy and Mathematics section of the New Jersey Assessment of Skills and Knowledge (NJ ASK) 6, 7, and 8. Additionally, this study aims to examine specific models, including the independent variables of inclusive setting, non-inclusive setting, gender, race, student attendance, and eligibility for free and reduced lunch that, paired with placement in an inclusive/non-inclusive setting, may cause an effect on the dependent variable of student achievement on the NJ ASK

6-8. As emphasis continues to be placed on the outcomes of the NJ ASK scores to determine school accountability, district administrators must consider the needs of all students. With that said, this study aims to produce research-based evidence to continue the development of policy and create the opportunity for educated, data-driven decisions that benefit all students. As a result, school professionals will be able to choose an instructional program that meets individual students needs and that will maximize both learning and student achievement.

Inclusion Variables

Results of this study suggest that the variable of inclusion was a statistically significant variable that influenced student performance in both language arts and mathematics in School A. In the areas of language arts and mathematics, it appears that non-disabled students placed in an inclusion classroom are scoring lower than non-disabled peers placed in a general education classroom. However, caution should be exercised with this statement because the inclusion variable raises some questions about its validity to solely measure all the components of inclusion. Instead, it is safer to conclude that the structural makeup of inclusion in this urban district raises some cause for concern, as it does not align with current research regarding academic homogeneous grouping.

Given the limitations of this study, it cannot be overlooked that the variable of inclusion in School A could be masked by the academic makeup of the general education students being placed within the inclusive setting. Examination of the average NJ ASK scores for the general education population both in and not in inclusive settings (See Appendix C) has shown that there is approximately a 20 point difference in the mean scores of the general education students

placed in and not in inclusive settings. Review of the literature has shown that the inclusion model in the schools violated the 10-30-60 rule and ability-grouping research.

At a time when accountability is mandated for all students and all students need to achieve proficiency, these findings present a bit of a concern for the individual schools and the district as a whole. It is apparent that in School A the structure of the inclusion model is not successful in terms of improving student achievement on one state-mandated test. Research findings by Slavin suggest that an inclusion model should be balanced with the 10-30-60 model, meaning that 10% of the students in the classroom are children with special needs or low academic achievers, 30% must be average to above average achievers, and 60% must be high achievers; otherwise, the mixed-ability grouping model has academic consequences and even a negative influence on achievement.

With that said, more of an effort must be made to ensure that practices such as inclusion are aligned with research-based conclusions regarding inclusive classroom makeup and mixed-ability grouping. After all, it is a common misconception that having mixed ability grouping in classrooms would lower expectations and standards for the non-disabled students and, as a result, school leaders and administrators attempted to segregate individuals with disabilities into self-contained or resource room settings. Research has shown that "every means of grouping students by ability or performance level has drawbacks that might be serious enough to offset any advantages" (Slavin, 1988, p. 72). However, homogeneous ability grouping has demonstrated a history of helping no students (Mosteller et al., 1996, p. 798). "Teachers to administrators have typically tried to reduce variability by assigning students to classes based on some indicator" whether it be gender, ability level, or a variety of test scores (Zaharias, Achillies,

& Cain, 1995, p. 8). Unfortunately, this practice is the opposite intention of inclusive practices. Research has demonstrated over time that ability grouping may stigmatize low achievers, put them in classes or groups for which teachers have low expectations, or lead to the creation of academic elites (Slavin, 1988). Grouping might doom children who are not in top tracks to second-class instruction and ultimately deprive students of the examples and stimulation provided by heterogeneous classes with theory and research support (Slavin, 1988; Zaharias, Achillies, & Cain, 1995).

Social Learning Theory

Bandura's social learning theory plays a large part in educating students in inclusive education settings and would explain why homogeneous groups of low achieving general education students are not reaching their academic potential in this study and why research supports the 10-30-60 structure of inclusive settings. Social learning theory incorporates aspects of behavioral (assumes that people's environment causes people to behave in certain ways) and cognitive learning (psychological factors are important for influencing how one behaves). This theory is a combination of environmental and psychological factors that establishes three requirements on how people learn and model behavior: retention (remembering what one observed), reproduction (ability to reproduce the behavior) and motivation (reason to want to adopt the behavior).

One of the first and most important implications of social learning theory is that *students often learn a great deal simply by observing others (*Ormrod, 2008, p. 145). The problem laced within the structure of the inclusion model used in School A is the lack of quality academic models. The implication is that students learn from watching their age appropriate peers.

Without peers to exhibit and model appropriate behaviors, children would be more likely to duplicate inappropriate behaviors that are displayed by other students. Observing appropriate behaviors by means of age-appropriate models is an integral piece of justifying inclusion as well as developing appropriate learning behaviors.

Without the high achieving 60% of the 10-30-60 inclusive structure, students would not have the necessary academic models, which is another educational implication for social learning theory. The regular education students enrolled in the inclusion environment become the models "providing an alternative to shaping new teaching behaviors" (Ormrod, 2008, p. 67) and clearly there is a lack of quality academic role models in the inclusive classrooms in School A. Students can serve as informal models, showing students proper behaviors in various settings, or they can be used as formal models or examples. This idea is one of the four "essential conditions that must exist in order to promote effective learning" (Ormrod, 2008, p. 145).

Additionally, the findings of this study directly align with production function theory, which is defined as describing the relationship between school inputs and student outcomes (Greenwald, Hedges & Laine, 1996). As stated by Caldas (1993), "The factors that affect student achievement over which school officials have little or no control have been termed *input factors*, whereas those factors over which school officials do have control have been termed *change or process factors*" (p. 224). In other words, gender, socio-economic status, and student attendance would be considered input factors because they comprise demographic and socioeconomic characteristics that can have an influence on student achievement but cannot be altered easily or in their entirety by school officials (Jencks, Smith, Ackland, Bane, Cohen, Gintis, Heyns and Michelson, 1972). Other variables known as "process variables" involved in this study which

can be changed or altered include the following: placement in an inclusive setting, and peer influence.

Inclusion as a Process

Inclusion is a process variable that is considered an input variable that can be changed but, if not done according to research-based practices, will have a negative influence on academic achievement, in this case for the general education student. This study supports previous research findings, including the following: the 10, 30, 60 rule, ability grouping having a negative impact on low performing students, the need for prior achievement to be examined to determine class makeup so that the average test scores in the inclusion class are not much different than the average score of the general education class.

The results rendered from School B found that pre-achievement and race were the strongest predictors of non-disabled student achievement as measured by the NJ ASK in both mathematics and language arts, both of which are input factors and not able to be altered. In other words, input variables had a more significant influence on student achievement, as inclusion was not found to be a significant variable in School B in the areas of mathematics or language arts.

It is important to note that in the analysis of School B, the previous year's test scores (Math 2010 and LA 2010) presented a situation where the partial correlation coefficient actually decreased the percentage of the variance accounted for. In examination of the hierarchical regression, Model 3, the Beta for Math 2010 was .811 (66% of the variance), but the partial coefficient was actually less, .783 (61% of the variance). Thus the partial coefficient is actually bringing the percentage of the variance accounted for from 66% to 61%, which means that if

Math 2010 (previous test scores) could possibly account for 5% less of the variance, then other variables could perhaps account for more of the variance. This is one possibility as to why the other variables in School B mathematics were not found to be statistically significant. Math 2010 was acting as a suppressor variable.

The same situation occurred with LA 2010. In examination of the hierarchical regression, Model 3, the Beta for LA 2010 was .688 (47.3% of the variance), but the partial coefficient was actually less, .627 (45% of the variance). So the partial coefficient is actually bringing the percentage of the variance accounted for from 47.3% to 45%, which means that if LA 2010 (previous test scores) could possibly account for approximately 2% less of the variance, then other variables could perhaps account for more of the variance. This is one possibility as to why the other variables in School B language arts, were not found to be statistically significant. LA 2010 was acting as a suppressor variable. Both of these statistical findings present the possibility that the inclusion variable could be suppressed and not show its true strength in this study.

Although the possibility exists that other variables could be suppressed by prior achievement, school administrators need to consider prior achievement scores when grouping students. Whether it is an inclusion classroom or a non-inclusive setting, it is important to consider the academic level of students being placed in any classroom because it is clear that prior achievement scores play a large role in determining academic outcomes on standardized assessments.

Recommendations for Policy and Practice

The findings of this study could be used to assist local teachers and administrators as they attempt to develop effective inclusion programs and address the issues of student achievement, which may ultimately affect proficiency goals in mathematics and language arts.

This study recommends that school administrators need to ensure that the 10-30-60 rule is not violated when designing inclusive classrooms. In order for an inclusive classroom to be successful, there needs to be quality academic role models (60%) available within the inclusive classroom, an even distribution of general education students in both inclusive and non-inclusive classrooms, and previous achievement needs to be considered and examined when making final decisions on the classrooms in which students are placed.

The challenge for school administrators when creating an inclusion class can also be financial, as the fewer students with disabilities in a classroom, the more inclusion classrooms are necessary in each school. As a result, more special education teachers are needed to provide support in additional inclusion classes. Although Special Education Code 6A:14-4.6 allows a maximum of eight special education students per inclusion classroom, this actually violates the 10-30-60 rule. This implies that administrators need to consider putting fewer students with disabilities in each inclusion class when possible even though the law allows up to eight students. Lawmakers should revise the law in order to discourage the practice of putting eight students with disabilities in each inclusion classroom; having eight special education students in an inclusion class violates the 10-30-60 rule. Following these basic rules will solidify the structural aspects of inclusion and support research based practice on classroom make up and ability grouping.

Additionally, this study rendered inclusion as a statistically significant variable that influenced the academic performance of general education students in both language arts and math. The research points to a few areas of concern that need to be examined when a school implements inclusion. First, what is the academic makeup of the students being placed in the inclusion classroom? Is the administrator creating the classroom roster following the research based 10-30-60 model and not flooding the classroom with all lower level learners as done in this data set? Are the school administrators grouping students according to a specific characteristic and if so, is this grouping going to influence student achievement? What is the attitude of the teacher, or teachers, toward inclusion? Do the professionals employ all the aspects of academic optimism in the school and inclusion classroom? These are the most important areas that need to be considered when a school has inclusion classrooms; otherwise, inclusion can have a negative influence on the academic achievement of general education students.

General Policy Recommendations

Federal policies are not one-size-fits-all; each district and school needs to examine each policy in order to determine how it works/aligns with each school. This study examined two schools in the same district that have different populations; therefore, the results have different outcomes. The results of this study find that between two schools in the same school district, the results of inclusive placement on non-disabled students' achievement is inconsistent and requires change and future investigation. The underlying problem with federal policies and mandates is that they are often designed as one-size-fits-all educational innovations, which have been found not to work because they ignore contextual factors that determine an intervention's efficacy in a particular local situation (Clarke & Dede, 2006, p. 27).

> When interventions [such as inclusion] are mandated, consideration must be given to variables within the intervention's setting that represent important conditions for success, and summarizing the extent to which the influence of the intervention is attenuated by variation in them can provide prospective adopters of the innovation a better sense of what level of effectiveness they are likely to enjoy in their own particular circumstances (Clarke & Dede, 2006, p. 28).

Policies developed at the federal and state level often operate with the notion that the same treatment should work with all students with only minor tailoring to the intervention. For example, the No Child Left Behind legislation requires all students to take standardized assessments in order to determine proficiency. This was a one-size-fits-all policy created by members of the federal government to increase the accountability of schoolteachers and administrators. With the creation of this policy, little consideration was given for formal adaptations to the culture, language, context, or special needs of the individuals who will be subject to this intervention.

Administrators are presented with the challenge of having to "fit" existing policies such as inclusion into their existing schools of students with little guidance or consideration for the need for adaptations to fit their population. The results of this study suggest that the same intervention of inclusion renders mixed results and requires change and adaptation in order for it to produce positive academic results. Inputs and processes matter. Research by Bernal (2009) and Clarke and Dede (2006) produced results that show that the one-size-fits-all model does not succeed because one pedagogical strategy that is successful in one particular classroom setting

with one particular group of students frequently will not succeed in a different classroom with different students.

Researchers found that there is a substantial influence of contextual variables (race, SES, prior achievement, teacher attitude/preparation) in shaping the desirability, practicality, and effectiveness of educational interventions. They concluded that identifying variables within the intervention's setting that represent important conditions for success and summarizing the extent to which the influence of the intervention is "attenuated by variation in such variables can provide prospective adopters of the innovation a better sense of the level of effectiveness they are likely to enjoy in their own particular circumstances (Clark & Dede, 2006, p. 29)." Additionally, these analytical thoughts can help designers of educational interventions improve the flexibility of their innovations by developing "hybrid" versions optimized for desired results in a variety of diverse settings. Given the results of this study, inclusion is not a successful one-size-fits-all intervention, as demonstrated by the two middle schools producing different results regarding the influence of inclusion. School principals and building level administrators need to use research-based methods to create the foundation for inclusive settings, examine inclusion classrooms on an annual basis, and make changes based on the context and climate of the school which he or she oversees.

Recommendation for Policy and Practice

The findings from this study must be shared with school leaders and administrators in order to address the issues of inclusion and the variety of results found within one urban school district. Inclusion, an education intervention designed to provide a quality education for students with disabilities, is not a "one size fits all" intervention and must continue to be examined to

develop those “hybrid” versions that fit each individual school; after all, inclusion does not only influence students with disabilities, it influences the general education students in the classroom as well.

Researchers must begin to expand on these findings by assessing the influence of inclusion on both disabled and non-disabled students in other curricular areas such as science and social studies to see if similar findings occur. Doing so has the potential to provide more evidence on the efficacy of inclusion programs in each school. It is also essential that school administrators begin to collect longitudinal data to track cohorts of students across multiple settings and years in an effort to assess the effectiveness of the inclusion programs offered in their school because as this study found, the structure of inclusion varies even between schools in the same district.

Finally, the results from this study add to the existing body of literature regarding the influence of the inclusive classroom on the academic achievement of non-disabled students. Although this study provided mixed results, it does provide additional insight in that inclusion needs to be examined by each individual school. The evidence collected from this study suggests that federal, state, and local agencies should reconsider the mandate of inclusion and any education intervention that aims to be one-size-fits-all. In some cases one intervention will be the navigation course of a student's entire future learning experience.

Recommendations for future practice of inclusion include school leaders examining the makeup of the students placed in the inclusion classroom, ensuring that there is a strategically placed diversity of learners in the classroom, examining the attitude and academic optimism of the teachers that are placed within the inclusion classroom, and being sure to examine NJ ASK

test scores in order to determine the success of inclusion. Ultimately, desired results should indicate that inclusion has no influence on the academic achievement of non-disabled students placed in an inclusion classroom.

Recommendations for Future Research

Although this study provided empirical evidence adding to the existing body of research on the influence of inclusion on general education students, it is not possible that a single explanatory study could provide all the answers to this research question. Therefore, it is important to conduct future research in the area of inclusion in order to determine how it should be implemented and examined in order to promote student achievement for all.

1. Recreate this study with a large enough sample in another urban, suburban, or rural school district and, in addition to the statistical analyses conducted in this study, include an analysis examining individual classrooms.

2. Conduct a longitudinal study with a large population, separating the general education students in inclusion by grades to compare and contrast the findings to this study to investigate whether specific grade levels have an influence on general education students placed in inclusion.

3. Design a study to examine teacher attitudes towards inclusion. It is possible that teacher attitude and expectations have an influence on the general education students' academic achievement when they are placed in the inclusion classroom.

4. The data and the findings of this study do suggest that further analysis using logistic regression could provide an odds ratio for the probability of passing the NJ ASK based on a student's placement in general education or inclusion classes.

The results of this study, the continued mandate of inclusion in conjunction with other student variables that cannot be altered (i.e., student's gender, SES. attendance record), and the required use of test results from state mandated assessments suggest that further study in the area of inclusion is necessary for the academic success for all students.

References

Affleck, J. Q., Madge, S., Adams, A., & Lowenbraun, S. (1988). Integrated classroom versus resource model: Academic viability and effectiveness. *Exceptional Children, 54*(4), 339-348.

Americans with Disabilities Act, 42 U.S.C. 12101. (1990).

Andrews, J. E,, Carnine, D. W., & Coutinho, M. J. (2000, September/October). Bridging the special education divide. *Remedial and Special Education, 21*(5), 258-267.

Apling, R. N., & Jones, N. L. (2005, January). *Individuals with disabilities education act (IDEA): Analysis of changes made by P.L. 108-446. Washington, DC:* CRS Report for Congress.

Apling, R., & Jones, N. L. (2002, January). *The individuals with disabilities education act (IDEA): Overview of major provisions*. Washington, DC: CRS Report for Congress.

Averich, H. A., Carroll, S. J., Donaldson, T. S., Kiesling, H. J., & Pincus, J. A. (1972). *How effective is schooling*? Santa Monica, CA: The Rand Corporation.

Baker, J. M., & Zigmond, N. (1995, Summer). The meaning and practice of inclusion for students with learning disabilities: Themes and implications from the five cases. *Journal of Special Education, 29*(2), 163-180.

Baker, E. T., Wang, M. C., & Walberg, H. J. (1994). The effects of inclusion on learning. *Educational Leadership, 52*(4): 33-35.

Baker, D. P., & Jones, D. P. (1993). Creating gender equality: Cross- national gender stratification and mathematical performance. *Sociology of Education, 66*, 91–103.

Baines, L., & Baines, C. (1994). Mainstreaming: one school's reality. *Phi Delta Kappan,*

76(1), 39-48.

Bennett, T., Deluca, D., & Bruns, D. (1997, Fall). Putting inclusion into practice: Perspectives of teachers and parents. *Exceptional Children, 64(*1), 115-142.

Bollen, K. A., Glanville, J. L., & Stecklov, G. (2001). Socioeconomic status and class in studies of fertility and health in developing countries. *Annual Review of Sociology*, 153-185.

Borland, M. V., & Howsen, R. M. (1998, March/April). Effect of student attendance on performance: Comment on Lamdin. *The Journal of Educational Research, 91*(4), 195-197.

Brown v. Board of Education. 347 U.S. 483 (1954).

Burke, M. A., & Sass, T. R. (2008, June). *Classroom peer effects and student achievement* (Working paper No. 18). Washington DC: National Center for Longitudinal Data in Education Research, Urban Institute.

Burris, C. C., Heubert, J. P., & Levin, H. M. (2006, Spring). Accelerating mathematics achievement using heterogeneous grouping. *American Educational Research Journal, 43*(1), 105-136.

Caldas, S. J. (1993, March/April). Reexamination of input and process factor effects on public school achievement. *Journal of Educational Research, 86*(4), 206-224.

Cawley, J., Hayden, S., Cade, E., & Baker-Kroczynski, S. (2002). Including students with disabilities into the general education science classroom. *Exceptional Children, 68*(4), 423-435.

Chen, C., & Stevenson, H. W. (1995). Motivation and mathematics achievement: A comparative study of Asian-American, Caucasian-American, and East Asian high school students. *Child Development, 66*, 1215-1234.

Civil Rights Act of 1964 (P.L. 88-352).

Clark, J., Dede, C., Ketelhut, D. J., Nelson, B., & Bowman, C. (2006). A design-based research strategy to promote scalability for educational innovations. *Education Technology, 46*(3), 27-36.

Clyde K. v. Puyallup School District, 35 F. 3d 1396 (9th Cir., 1994).

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. ((1966*). Equality of educational opportunity* [Summary Report]. Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education.

Coleman, J. S., Kelly, S. D., & Moore, J. A. (1975). Recent trends in school integration. *Educational Researcher*, 3-12.

Daniel, L. G., & King, D. A. (1997). Influence of inclusion education on academic achievement, Student behavior and self-esteem, and parental attitudes. *The Journal of EducationalResearch, 91*(2), 67-80.

Daniel R. R. v. State Board of Education, 874 F. 2d 1036 (5th Cir. 1989).

De Klerk, G. (2008). Classical test theory (CTT). In M. Born, C. D. Foxcroft, & R. Butter (Eds.), *Online readings in testing and assessment (pp. 1-9).* Boston, MA: International Test Commission. Retrieved June 8, 2010 from http:/www.intestcom.orgPublicationsORTA.php.

Dietz, S., Jennings, J., Retner, D. S., & Cober, N. (2010, December). *State high school tests: Exit exams and other assessments* (Research Report No. 3). Washington, DC: Center for Education Policy.

Dornbusch, S. M., Ritter, P. L., & Steinberg, L. (1991). Community influences on the relation of family statuses to adolescent school performance: Differences between African-Americans and non-Hispanic Whites. *American Journal of Education*, 543-567.

Drukker, M., Kaplan, C., Feron, F., & Van Os, J. (2003). Children' health-related quality of life, neighborhood socioeconomic deprivation and social capital: A contextual analysis. *Social Science Medicine, 57*, 825-841.

Duncan, O. D., Featherman, D. L., & Duncan, B. ((1972). *Socioeconomic background and achievement.* New York, NY: Seminar Press.

Education for All Handicapped Children Act of 1975 (P.L. 94-142).

Elementary and Secondary Act of 1965 (P.L. 89-10).

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*(1), 103-127.

Fletcher, J. (2010). Spillover effects of classmates with emotional problems on test scores in early elementary school. *Journal of Policy Analysis and Management, 29*(1), 69-83.

Florian, L., Rouse, M., Black-Hawkins, K., & Jull, S. (2004). What can national data sets tell us about inclusion and pupil achievement? *British Journal of Special Education, 31*(3), 115-120.

Ford, J., & Sutphen, R. (1996, April). Early intervention to improve attendance in elementary school for at-risk children: A pilot program. *Social Work in Education, 18*(2), 95-102.

Fox, N. E., & Ysseldyke, J. E. (1997). Implementing inclusion at the middle school level. *Exceptional Children, 64*(1), 81-98.

Fraenkel, J. R., & Wallen, N. E. (2006). *How to design and evaluate research in education* (6th ed.). Boston, MA: McGraw Hill.

Freeman, S. (2012). Original position. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Spring ed., pp. 1-37). Retrieved from http://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=original-position

Fryxell, D., & Kennedy, C. H. (1995). Placement along the continuum of services and its influence on students' social relationships. *Journal of the Association for Persons with Severe Handicaps, 20*, 259-269.

Gamoran, A., & Long, D. A. (in press). *Equality of educational opportunity: A 40-year retrospective*. Madison, WI: University of Wisconsin.

Gandhi, A. G. (2007, March). Context matters: Exploring relations between inclusion and reading achievement of students without disabilities. *International Journal of Disability, Development and Education, 54(*1), 91-112.

Gelb, S. A. (1995). Darwin's use of intellectual disability. *Disabilities Studies Quarterly, 28*(2), 1-25.

Goals 2000: Education America Act. (H.R. 1804, 103d Cong. (1993).

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996, Fall). The effect of school resources on student achievement. *Review of Educational Research, 66*(3), 361-396.

Grissmer, D., Flanagan, A., & Williams, S. (1998). Why did the Black-White score gap narrow in the 1970s and 1980s? In C. Jencks & M. Phillips (Eds.), The Black-White test score gap (pp. 182-228). Washington, DC: The Brookings Institute.

Gunzelman, B., & Connell, D. (2006). The new gender gap: Social, psychological, neuro-biological, and educational perspectives. *Educational Horizons, 84*(2), 94-101.

Gurian, M., & Stevens, K. (2006). How are the boys doing? How boys learn. *Educational Horizons, 84*(2), 87-93.

Hardman, M. L., Drew, C. J., & Egan, M. W. (2002). *Human Exceptionality* (7th ed.). Boston, MA: Allyn & Bacon.

Harwell, M., & LeBeau, B. (2010). Student eligibility for free lunch as an SES measure in education research. *Educational Researcher, 39*(2), 120-131.

Hoffer, T. B. (1992, Fall). Middle school ability grouping and student achievement in science and mathematics. *Educational Evaluation and Policy Analysis, 14*(3), 205-227.

Hoy, W. K., Tarter, C. J., & Hoy, A. W. (n.d.). Academic optimism of schools and student achievement. Manuscript in preparation, School of Educational Policy and Leadership, Ohio State University, Columbus, Ohio.

Huber, K. D., Rosenfeld, J. G., & Fiorello, C. A. (2001). The differential influence of inclusion and inclusive practices on high, average, and low achieving general education students. *Psychology in the Schools, 38*(6), 497-504.

Hunt, P., & Goetz, L. (1997). Research on inclusive educational programs, practices, and outcomes for students with severe disabilities. *Journal of Special Education, 31*(1), 3-29.

Hunt, P., Staub, D., Alwell, M., & Goetz, L. (1994) Evaluating the effects of placement of students with severe disabilities in general education versus special classes. *Journal of the Association for Persons with Severe Handicaps, 19* (3),

200-14.

Hyde, J.S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A., & Williams, C. (2008). Gender similarities characterize math performance. *Science, 321*, 494-495.

IDEA Amendments of 1997 (P.L. 105-17).

IDEIA Amendments of 2004 (P.L. 108-446).

Idol, L. (2006, March/April). Toward inclusion of students in special education in general education. *Remedial and Special Education, 27*(2), 77-94.

Improving America's schools: A newsletter on issues in school reform. (1996). [Fact sheet]. Retrieved from http://www2.ed.gov/pubs/IASA/newsletters/standards/index.html

Individuals with Disabilities Education Act (IDEA) of 1990 (P.L. 101-476).

Jenkins, J. R., Jewell, M., Leicester, M., Jenkins, L., & Troutner, N. M. (1991). Development of a school building model for educating children with handicaps and at-risk students in general education classrooms. *Journal of Learning Disabilities, 24*(5), 311-320.

Kain, J. F., & Singleton, K. (1996, May/June). Equality of educational opportunity revisited. *New England Economic Review*, 87-108.

Kalambouka, A., Farrell, P., Dyson, A., & Kaplan, I. (2007). The influence of placing students with special education needs in mainstream schools on the achievement of their peers. *Educational Research, 49*(4), 365-382.

Kaufman, J. M. & Hallahan, D. P. (1995). *The illusion of full inclusion: A comprehensive*

critique of a special education bandwagon. Austin, TX: Pro-Ed.

Klingner, J. K., Vaughn, S., Hughes, M. T., Schumm, J. S., & Elbaum, B. (1998). Outcomes for students with and without learning disabilities in inclusive classrooms. *Learning Disabilities Research and Practice, 13*(3), 153-161.

Kraut, R. (2012). Plato. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy.* (Summer ed., pp. 1-22). Retrieved from http://plato.stanford.edu.cgi-bin/encyclopedia/archinfo.cgi?entry=plato

Lamdin, D. J. (1996, January/February). Evidence of student attendance as an independent variable in education production functions. *The Journal of Educational Research, 89*(3), 155-162.

La Morte, M. W. (2008). *School Law.* Boston, MA: Pearson.

Latham, A. S. (1997-1998). Gender Differences on Assessments. *Educational Leadership, 55*(4), 88-89.

Leech, N. L., Barrett, K. C., and Morgan, G. A. (2011). IBM SPSS for intermediate statistics: Use and interpretation (4th ed.). New York, NY: Routledge-Taylor & Francis Group.

Lewis, A. (2002). Accessing through research interviews the views of children with difficulties in learning. *Support for Learning, 17*, 11-116. d.o.i.:10.1111/1467-9604.00248

Luzader, C. E. (1995). Inclusion. *LINK*, 14(1), 2-30.

Marks, G. N. (2008). Accounting for the gender gaps in student performance in reading and mathematics: Evidence from 31 countries. *Oxford Review of Education, 34*(1), 89-109.

Martin, E. W., Martin, R., & Terman, D. L. (1996, Spring). The legislative and litigation history of special education. *The Future of Special Education for Students with Disabilities, 6*(1), 25-39.

Mastropieri, M. A., & Scruggs, T. E. (2001). Promoting inclusion in secondary classrooms. *Learning Disabilities Quarterly, 24*, 265-274.

McDonnell, J., Thorsen, N., Disher, S., Mathot-Buckner, C., Mendel, J., & Lavinia, R. (2003). The achievements of students with developmental disabilities and their peers without disabilities in inclusive settings: An exploratory study. *Education and Treatment of Children, 26*(3), 224-236.

Mills v. Board of Education of the District of Columbia, 348 F. Supp. 866 (DC, 1972).

Mosteller, F., Light, R. J., & Sachs, J. A. (1996, Winter). Sustained inquiry in education: Lessons from skill grouping and class size. *Harvard Educational Review, 66*(4), 797-828.

Mullis, I. V., Martin, M. O., Fierros, E. G., Goldberg, A. L., & Stemler, S. E. (2000). Gender differences in achievement. Boston, MA: Boston College, Lynch School of Education.

National study of inclusive education. (1994). New York, NY: National Center on Education Restructuring and Inclusion.

National study on inclusion: Overview and summary report. (1995). New York, NY: National Center on Educational Restructuring and Inclusion.

New Jersey Department of Education. (2010a). *Executive order No.14*. Retrieved from http://www.state.nj.us/education/news/2010/0209rc.htm

New Jersey Department of Education. (2010b). *New Jersey department of education releases 2009 school report curds*. Retrieved from http://www.state.nj.us/education/news/2010/0209rc.htm

New Jersey Department of Education (2009). *New Jersey assessment of skills and knowledge: 2008 technical report grades 5-8*. Retrieved February 1, 2012 from http://www.state.nj.us/education/assessment/ms/5-8/tech/2008techreport.pdf

New Jersey Department of Education (2006a). *District factor groups for school districts.* Retrieved from http://www.state.nj.us/education/finance/sf/dfgdesc.shtml

New Jersey Department of Education. (2006b). *New Jersey assessment of knowledge and skills.* Retrieved from http://www.state.nj.us/education/assessment/

New Jersey Department of Education (2006c). *New Jersey core curriculum content standards.* Retrieved from http://www.state.nj.us/education/cccs/

No Child Left Behind Act of 2001 (P.L. 107-110).

Orlich, D. C. (2006). *School reform*. Baltimore, MD: Publish America.

Ormrod, J. E. (2007). *Human learning* (5th ed.). Saddle River, NJ: Prentice Hall.

Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory into Practice, 41*(2), 116-125.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction.* Fort Worth, TX: Harcourt Brace College Publishers.

Peltier, G. L. (1997, Summer). The effect of inclusion on non-disabled children: A review of research. *Contemporary Education 68*(4), 234-238.

Pennsylvania Association for Retarded Children v. Pennsylvania, 343 F. Supp. 279 (E.D. PA, 1972).

Perie. M., Moran, R., & Lutkus, A. D. (2005). NAEP 2004 trends in academic progress: Three decades of student performance in reading and mathematics (NCES 2005-464). Washington, DC: Government Printing Office.

Publication manual of the American Psychological Association (6th ed.). (2010). Washington DC: Author.

Raymond, M. E., & Hanushek, E. A. (2003, Summer). High-stakes research. *Educational Measurement: Issues and Practice, 24*(4), 32-39. Retrieved from http://www.educationnext.org/

Rea, P., McLaughlin, V., & Thomas, C. (2002). Outcomes for students with disabilities in inclusive and pullout programs. *Exceptional Children, 68,* 203-208.

Robins, L. N., & Ratcliff, K. S. (1978). *Public health service: Long range outcomes associated with school truancy* (Report No. 152893). Rockville, MD: Author.

Roby, D. E. (2004). Research on school attendance and school achievement: A study of Ohio schools. *Educational Research Quarterly, 28*(1), 4-14.

Rubinfeld, D. L. (2000). Reference guide on multiple regression. In E. Beecher-Monas (Ed.), *Reference manual on scientific evidence* (pp. 179-227). Retrieved from http://www.law.berkeley.edu/faculty/rubinfeld/Profile/publications/Rubinfeld_Multiple %20Regressionpdf

Sacramento City Unified School District Board of Education v. Rachel Holland, 14 F. 3d 1398 (9th Cir. 1994).

Saint-Laurent, L., Dionne, J., Giasson, J., & Royer, E. (1998). Academic achievement effects of an in-class service model on students without disabilities. *Exceptional Children 64*, 15.

Salend, S. J., & Duhaney, L. G. (1999). The influence of inclusion on students with and without disabilities and their educators. *Remedial and Special Education, 20*(2), 114-127.

Sharpe, M. N., York, J. L., & Knight, J. (1994, Sept.). The effects of inclusion on the academic performance of classmates without disabilities. *Remedial and Special Education, 15*(5), 281-287.

Shattuck, R. (1981). *The forbidden experiment: The story of the wild boy of Aveyron*. New York, NY: Pocket.

Sheldon, S. B. (2007, May/June). Improving student attendance with school, family and community partnerships. *The Journal of Educational Research, 100*(5), 267-275.

Sirin, S. R. (2005, Fall). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417-453.

Slavin, R. E. (1988, September). Synthesis of research on grouping in elementary and secondary schools. *Educational Leadership*, 67-77.

Slavin, R. E. (1987). Grouping for instruction in the elementary school. *Educational Psychologist, 22*(2), 109-127.

Slavin, R. E., & Lake, C. (2008, September). Effective programs in elementary mathematics: A best evidence synthesis. *Review of Educational Research, 78*(3), 427-515.

Staub, D., & Peck C.A. (1995). What are the outcomes for non-disabled students?

Educational Leadership, 52(4), 36-40.

Sternberg, R. J., & Williams, W. M. (2002). *Educational Psychology.* Boston: Allyn & Bacon.

Tienken. C. H. (2008a). A descriptive study of the technical characteristics of the results from New Jersey's 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). The characteristics of state assessment results. *Academic Exchange Quarterly, 12*(3), 34-39.

Trabucco, M. (2011). *The influence of co-taught inclusion on the academic achievement of non-disabled students in mathematics as measured by the New Jersey assessment of skills and knowledge* (Unpublished doctoral dissertation). Seton Hall University, South Orange, NJ.

U.S. Department of Education. (1983). A nation at risk: The imperative for education reform. Retrieved from http://datacenter.spps.org./uploads/ SOTW_A_Nation_At_Risk_1983.pdf

Usher, A. (2011, December 15). *Annual yearly progress results 2010-2011. Washington, DC:* Center on Education Policy. Retrieved from http:///www.cep-dc.org/

Velde, F. (1999). Antoin E. Murphy's John Law: Economic theorist and policymaker. Journal of Political Economy, 107(1), 201-204. Retrieved from http://www.jstor.org/stable/ 10.1086/250058

Wilkins, J. L., Zembylas, M., & Travers, K. J. (2002). Investigating correlates of mathematics and science literacy in the final year of secondary school. In D. F. Robataille & A. E. Beaton (Eds.), *Secondary analysis of the TIMMS data* (pp. 291-316). Boston, MA: Kluwer.

Willingham, W. W., & Cole, N. S. (1997). *Gender and fair assessment.* Mahwah, NJ: Lawrence Erlbaum Associates.

Witte, R. S. & Whitte, J. S. (2007). *Statistics.* (8^{th} ed.), Hoboken, NJ: John Wiley & Sons.

Wright, P., & Wright, P. (2010). *Least Restrictive Environment and FAPE.* Retrieved November 20, 2011 from http://www.wrightslaw.com/advoc/articles/idea.lre.fape.htm.

Zaharias, J. B., Achilles, C. M., & Cain, V. A. (1995). The effect of random class assignment on elementary students' reading and mathematics achievement. *Research in the Schools, 2*(2), 7-14.

Appendix A

Chi Square Analysis

Chi Square Analysis for Race, Total Sample Population

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
School Code * Race	1280	100.0%	0	0.0%	1280	100.0%

School Code * Race Crosstabulation

Count

		Race		
		0	1	Total
School Code	0	111	518	629
	1	231	420	651
Total		342	938	1280

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	51.981[a]	1	.000		
Continuity Correction[b]	51.074	1	.000		
Likelihood Ratio	52.894	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	51.941	1	.000		
N of Valid Cases	1280				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 168.06.

b. Computed only for a 2x2 table

Chi Square Analysis for 2011 NJ ASK Math Scores, Total Sample Population

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
School Code * MAth 2011	1275	99.6%	5	0.4%	1280	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	130.189[a]	113	.128
Likelihood Ratio	148.270	113	.015
Linear-by-Linear Association	7.200	1	.007
N of Valid Cases	1275		

a. 115 cells (50.4%) have expected count less than 5. The minimum expected count is .49.

Chi Square Analysis for 2011 NJ ASK Language Arts Scores, Total Sample Population

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
School Code * LAL 2011	1271	99.3%	9	0.7%	1280	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	111.620[a]	101	.221
Likelihood Ratio	129.396	101	.030
Linear-by-Linear Association	.054	1	.816
N of Valid Cases	1271		

a. 107 cells (52.5%) have expected count less than 5. The minimum expected count is .49.

Chi Square Analysis for Economically Disadvantaged Students, Total Sample Population

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
School Code * Economically Disadvantaged	1280	100.0%	0	0.0%	1280	100.0%

School Code * Economically Disadvantaged Crosstabulation

Count

		Economically Disadvantaged		Total
		0	1	
School Code	0	203	426	629
	1	345	306	651
Total		548	732	1280

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	56.106[a]	1	.000		
Continuity Correction[b]	55.263	1	.000		
Likelihood Ratio	56.602	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	56.062	1	.000		
N of Valid Cases	1280				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 269.29.

b. Computed only for a 2x2 table

Chi Square Analysis for Students in Special Education, Total Sample Population

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
School Code * Spec Ed	1280	100.0%	0	0.0%	1280	100.0%

School Code * Spec Ed Crosstabulation

Count

		Spec Ed		
		0	1	Total
School Code	0	489	140	629
	1	536	115	651
Total		1025	255	1280

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.229[a]	1	.040		
Continuity Correction[b]	3.946	1	.047		
Likelihood Ratio	4.233	1	.040		
Fisher's Exact Test				.042	.023
Linear-by-Linear Association	4.226	1	.040		
N of Valid Cases	1280				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 125.31.

b. Computed only for a 2x2 table

Chi Square Analysis for Students Eligible for Free Lunch, Total Sample Population

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
School Code * Free Lunch Eligible	1280	100.0%	0	0.0%	1280	100.0%

School Code * Free Lunch Eligible Crosstabulation

Count

		Free Lunch Eligible		Total
		0	1	
School Code	0	308	321	629
	1	429	222	651
Total		737	543	1280

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	37.548[a]	1	.000		
Continuity Correction[b]	36.858	1	.000		
Likelihood Ratio	37.729	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	37.519	1	.000		
N of Valid Cases	1280				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 266.83.

b. Computed only for a 2x2 table

Appendix B

Multiple Correlation Analysis

School A Multiple Correlation Analysis NJ ASK LAL 2011

Correlations

		LAL 2011	INC LA	LAL 2010	Free Lunch Eligible	Gender	Race	Absent
Pearson Correlation	LAL 2011	1.000	-.580	.714	-.132	.120	-.167	-.031
	INC LA	-.580	1.000	-.433	.056	-.064	.105	-.022
	LAL 2010	.714	-.433	1.000	-.134	.081	-.162	-.021
	Free Lunch Eligible	-.132	.056	-.134	1.000	-.011	.192	.124
	Gender	.120	-.064	.081	-.011	1.000	.073	-.019
	Race	-.167	.105	-.162	.192	.073	1.000	-.037
	Absent	-.031	-.022	-.021	.124	-.019	-.037	1.000
Sig. (1-tailed)	LAL 2011	.	.000	.000	.002	.004	.000	.249
	INC LA	.000	.	.000	.111	.079	.010	.316
	LAL 2010	.000	.000	.	.002	.038	.000	.326
	Free Lunch Eligible	.002	.111	.002	.	.407	.000	.003
	Gender	.004	.079	.038	.407	.	.055	.336
	Race	.000	.010	.000	.000	.055	.	.208
	Absent	.249	.316	.326	.003	.336	.208	.
N	LAL 2011	486	486	486	486	486	486	486
	INC LA	486	486	486	486	486	486	486
	LAL 2010	486	486	486	486	486	486	486
	Free Lunch Eligible	486	486	486	486	486	486	486
	Gender	486	486	486	486	486	486	486
	Race	486	486	486	486	486	486	486
	Absent	486	486	486	486	486	486	486

School A Multiple Correlation Analysis NJ ASK Math 2011

Correlations

		Math 2011	Absent	Free Lunch Eligible	Gender	Race	INC Math	MAth 2010
Pearson Correlation	Math 2011	1.000	-.094	-.098	-.132	-.181	-.681	.797
	Absent	-.094	1.000	.124	-.019	-.037	.057	-.092
	Free Lunch Eligible	-.098	.124	1.000	-.011	.192	.095	-.132
	Gender	-.132	-.019	-.011	1.000	.073	.046	-.106
	Race	-.181	-.037	.192	.073	1.000	.103	-.225
	INC Math	-.681	.057	.095	.046	.103	1.000	-.595
	MAth 2010	.797	-.092	-.132	-.106	-.225	-.595	1.000
Sig. (1-tailed)	Math 2011	.	.019	.015	.002	.000	.000	.000
	Absent	.019	.	.003	.336	.208	.105	.022
	Free Lunch Eligible	.015	.003	.	.407	.000	.019	.002
	Gender	.002	.336	.407	.	.055	.154	.010
	Race	.000	.208	.000	.055	.	.012	.000
	INC Math	.000	.105	.019	.154	.012	.	.000
	MAth 2010	.000	.022	.002	.010	.000	.000	.
N	Math 2011	486	486	486	486	486	486	486
	Absent	486	486	486	486	486	486	486
	Free Lunch Eligible	486	486	486	486	486	486	486
	Gender	486	486	486	486	486	486	486
	Race	486	486	486	486	486	486	486
	INC Math	486	486	486	486	486	486	486
	MAth 2010	486	486	486	486	486	486	486

School B Multiple Correlation Analysis NJ ASK LAL 2011

Correlations

		LAL 2011	Gender	Race	ABSENT	Free Lunch Eligible	INCLAL	LAL2010
Pearson Correlation	LAL 2011	1.000	.139	-.237	.015	-.174	-.270	.721
	Gender	.139	1.000	.007	-.042	-.035	-.012	.165
	Race	-.237	.007	1.000	.045	.279	.162	-.197
	ABSENT	.015	-.042	.045	1.000	.057	.041	-.019
	Free Lunch Eligible	-.174	-.035	.279	.057	1.000	.067	-.156
	INCLAL	-.270	-.012	.162	.041	.067	1.000	-.338
	LAL2010	.721	.165	-.197	-.019	-.156	-.338	1.000
Sig. (1-tailed)	LAL 2011	.	.001	.000	.369	.000	.000	.000
	Gender	.001	.	.432	.168	.212	.395	.000
	Race	.000	.432	.	.150	.000	.000	.000
	ABSENT	.369	.168	.150	.	.093	.169	.329
	Free Lunch Eligible	.000	.212	.000	.093	.	.061	.000
	INCLAL	.000	.395	.000	.169	.061	.	.000
	LAL2010	.000	.000	.000	.329	.000	.000	.
N	LAL 2011	535	535	535	535	535	535	535
	Gender	535	535	535	535	535	535	535
	Race	535	535	535	535	535	535	535
	ABSENT	535	535	535	535	535	535	535
	Free Lunch Eligible	535	535	535	535	535	535	535
	INCLAL	535	535	535	535	535	535	535
	LAL2010	535	535	535	535	535	535	535

School B Multiple Correlation Analysis NJ ASK Math 2011

Correlations

		Math 2011	Gender	Race	ABSENT	INCMath	Free Lunch Eligible	Math 2010
Pearson Correlation	Math 2011	1.000	-.032	-.282	-.071	-.355	-.173	.826
	Gender	-.032	1.000	.007	-.042	.015	-.035	-.064
	Race	-.282	.007	1.000	.045	.171	.279	-.299
	ABSENT	-.071	-.042	.045	1.000	.081	.057	-.087
	INCMath	-.355	.015	.171	.081	1.000	.099	-.411
	Free Lunch Eligible	-.173	-.035	.279	.057	.099	1.000	-.207
	Math 2010	.826	-.064	-.299	-.087	-.411	-.207	1.000
Sig. (1-tailed)	Math 2011	.	.230	.000	.051	.000	.000	.000
	Gender	.230	.	.432	.168	.364	.212	.071
	Race	.000	.432	.	.150	.000	.000	.000
	ABSENT	.051	.168	.150	.	.030	.093	.022
	INCMath	.000	.364	.000	.030	.	.011	.000
	Free Lunch Eligible	.000	.212	.000	.093	.011	.	.000
	Math 2010	.000	.071	.000	.022	.000	.000	.
N	Math 2011	535	535	535	535	535	535	535
	Gender	535	535	535	535	535	535	535
	Race	535	535	535	535	535	535	535
	ABSENT	535	535	535	535	535	535	535
	INCMath	535	535	535	535	535	535	535
	Free Lunch Eligible	535	535	535	535	535	535	535
	Math 2010	535	535	535	535	535	535	535

Appendix C

School Mean for NJ ASK Scores by Group

School A: Mean NJ ASK 2011 Scores for General Education Students in Inclusion for LA

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
LAL 2011	164	147	218	182.21	11.834
Valid N (listwise)	164				

School A: Mean NJ ASK 2011 Scores for Students In Special Education in Inclusion for LA

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
LAL 2011	164	147	218	182.21	11.834
Valid N (listwise)	164				

School A: Mean NJ ASK 2011 Scores for General Education Students NOT in Inclusion for LA

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
LAL 2011	325	100	257	209.66	21.163
Valid N (listwise)	325				

School A: Mean NJ ASK 2011 Scores for General Education Students in Inclusion for Math

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Math 2011	157	128	220	178.42	16.573
Valid N (listwise)	157				

School A: Mean NJ ASK 2011 Scores for Students In Special Education in Inclusion for Math

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Math 2011	23	156	218	179.22	15.133
Valid N (listwise)	23				

School A: Mean NJ ASK 2011 Scores for General Education Students NOT in Inclusion for Math

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Math 2011	332	163	300	226.66	27.234
Valid N (listwise)	332				

School B: Mean NJ ASK 2011 Scores for General Education Students in Inclusion for LA

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
LAL 2011	91	152	224	186.05	16.304
Valid N (listwise)	91				

School B: Mean NJ ASK 2011 Scores for Students in Special Education in Inclusion for LA

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
LAL 2011	15	138	203	179.20	17.022
Valid N (listwise)	15				

School B: Mean NJ ASK 2011 Scores for General Education Students NOT in Inclusion for LA

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
LAL 2011	445	131	267	201.80	21.981
Valid N (listwise)	445				

School B: Mean NJ ASK 2011 Scores for General Education Students in Inclusion for Math

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Math 2011	77	144	256	186.58	21.356
Valid N (listwise)	77				

School B: Mean NJ ASK 2011 Scores for Students in Special Education in Inclusion for Math

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Math 2011	19	145	220	181.11	20.300
Valid N (listwise)	19				

School B: Mean NJ ASK 2011 Scores for General Education Students NOT in Inclusion for Math

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Math 2011	459	134	300	220.87	33.219
Valid N (listwise)	459				

Appendix D

Schools A and B Factorial Univariate Analysis of Variance

Schools A and B Univariate Analysis of Variance , Language Arts and Literacy

Univariate Analysis of Variance

[DataSet1] C:\Users\babogera\Desktop\CRobinson_9-10-12

Between-Subjects Factors

		N
inc LAL		895
	Y	271
School Code	MMS	613
	SMS	553

Descriptive Statistics

Dependent Variable:LAL 2011

inc LAL	School Code	Mean	Std. Deviation	N
	MMS	197.58	24.716	503
	SMS	202.17	25.682	392
	Total	199.59	25.232	895
Y	MMS	184.21	16.863	110
	SMS	181.74	11.785	161
	Total	182.74	14.091	271
Total	MMS	195.18	24.044	613
	SMS	196.22	24.367	553
	Total	195.68	24.193	1166

Levene's Test of Equality of Error Variances[a]

Dependent Variable:LAL 2011

F	df1	df2	Sig.
4.864	3	1162	.002

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + LAL2010 + incLAL + SchoolCode + incLAL * SchoolCode

Tests of Between-Subjects Effects

Dependent Variable:LAL 2011

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	393563.107[a]	4	98390.777	396.233	.000	.577	1584.933	1.000
Intercept	78487.034	1	78487.034	316.078	.000	.214	316.078	1.000
LAL2010	329458.943	1	329458.943	1326.777	.000	.533	1326.777	1.000
incLAL	6874.517	1	6874.517	27.685	.000	.023	27.685	1.000
SchoolCode	4.307	1	4.307	.017	.895	.000	.017	.052
incLAL * SchoolCode	4495.496	1	4495.496	18.104	.000	.015	18.104	.989
Error	288293.999	1161	248.315					
Total	45327251.000	1166						
Corrected Total	681857.105	1165						

a. R Squared = .577 (Adjusted R Squared = .576)

b. Computed using alpha = .05

Estimated Marginal Means

School Code

Estimates

Dependent Variable:LAL 2011

School Code	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
MMS	194.456[a]	.835	192.818	196.095
SMS	194.310[a]	.740	192.857	195.763

a. Covariates appearing in the model are evaluated at the following values: LAL2010 = 194.35.

Pairwise Comparisons

Dependent Variable:LAL 2011

(I) School Code	(J) School Code	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
MMS	SMS	.146	1.110	.895	-2.032	2.325
SMS	MMS	-.146	1.110	.895	-2.325	2.032

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

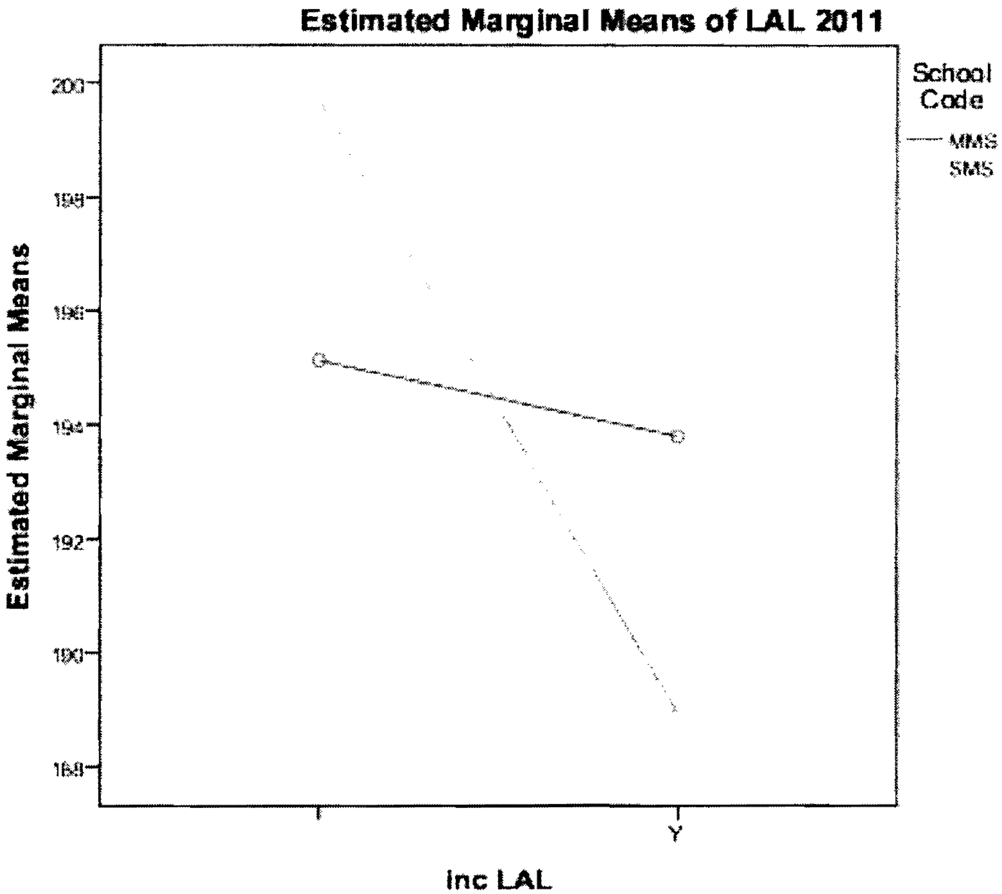
Univariate Tests

Dependent Variable:LAL 2011

	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Contrast	4.307	1	4.307	.017	.895	.000	.017	.052
Error	288293.999	1161	248.315					

The F tests the effect of School Code. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Computed using alpha = .05



Covariates appearing in the model are evaluated at the following values: LAL2010 = 194.

Schools A & B Univariate Analysis of Variance, Mathematics

Univariate Analysis of Variance

[DataSet1] C:\Users\babogera\Desktop\CRobinson_9-1

Between-Subjects Factors

		N
School Code	MMS	614
	SMS	558
Inc Math		928
	Y	244

Descriptive Statistics

Dependent Variable:Math2011

School Code	Inc Math	Mean	Std. Deviation	N
MMS		215.32	36.673	512
	Y	184.98	21.240	102
	Total	210.28	36.376	614
SMS		215.79	34.943	416
	Y	178.03	14.725	142
	Total	206.18	35.153	558
Total		215.53	35.889	928
	Y	180.93	18.031	244
	Total	208.33	35.842	1172

Levene's Test of Equality of Error Variances[a]

Dependent Variable:Math2011

F	df1	df2	Sig.
2.765	3	1168	.041

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Math2010 + SchoolCode + IncMath + SchoolCode * IncMath

Tests of Between-Subjects Effects

Dependent Variable:Math2011

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	987053.052[a]	4	246763.263	556.714	.000	.656	2226.855	1.000
Intercept	59066.667	1	59066.667	133.258	.000	.102	133.258	1.000
Math2010	752839.752	1	752839.752	1698.455	.000	.593	1698.455	1.000
SchoolCode	495.930	1	495.930	1.119	.290	.001	1.119	.185
IncMath	6790.506	1	6790.506	15.320	.000	.013	15.320	.974
SchoolCode * IncMath	4534.345	1	4534.345	10.230	.001	.009	10.230	.892
Error	517272.497	1167	443.250					
Total	52371270.000	1172						
Corrected Total	1504325.549	1171						

a. R Squared = .656 (Adjusted R Squared = .655)

b. Computed using alpha = .05

Estimated Marginal Means

School Code

Estimates

Dependent Variable:Math2011

School Code	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
MMS	207.478[a]	1.155	205.212	209.745
SMS	205.856[a]	1.046	203.804	207.908

a. Covariates appearing in the model are evaluated at the following values: Math2010 = 212.15.

Pairwise Comparisons

Dependent Variable:Math2011

(I) School Code	(J) School Code	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
MMS	SMS	1.622	1.533	.290	-1.387	4.630
SMS	MMS	-1.622	1.533	.290	-4.630	1.387

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

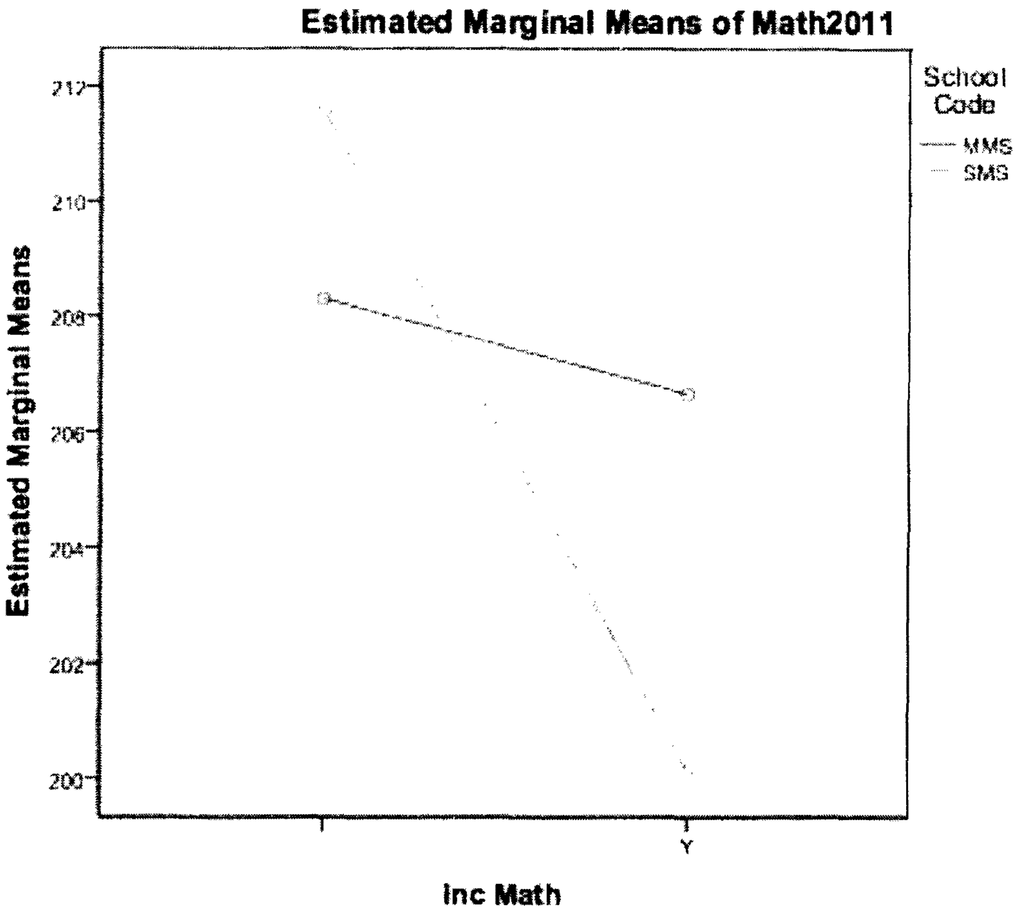
Univariate Tests

Dependent Variable:Math2011

	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Contrast	495.930	1	495.930	1.119	.290	.001	1.119	.185
Error	517272.497	1167	443.250					

The F tests the effect of School Code. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Computed using alpha = .05



Covariates appearing in the model are evaluated at the following values: Math2010 = 212.