

THE IMPACT OF SELECTED INITIATIVES ON THE READING CRITERION
REFERENCED COMPETENCY TEST SCORES OF AFRICAN-AMERICAN AND
DISADVANTAGED STUDENTS IN GRADES 3, 5, AND 8

by

Eric Angus Fuller

Liberty University

A Dissertation Presented in Partial Fulfillment
Of the Requirements for the
Doctor of Education

Liberty University

July, 2013

THE IMPACT OF SELECTED INITIATIVES ON THE READING CRITERION
REFERENCED COMPETENCY TEST SCORES OF AFRICAN AMERICAN AND
DISADVANTAGED STUDENTS IN GRADES 3, 5, AND 8

By Eric Angus Fuller

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University, Lynchburg, Virginia

July, 2013

APPROVED BY:

Dr. Rollen Fowler, Ph.D. , Chair

Dr. Amy T. McLemore, Ed.D. , Committee Member

Dr. Chanelle Sweet-White, Ed.D. , Committee Member

Dr. Scott B. Watson, Ph.D. , Associate Dean, Advanced Programs

THE IMPACT OF SELECTED INITIATIVES ON THE READING CRITERION
REFERENCED COMPETENCY TEST SCORES OF AFRICAN-AMERICAN AND
DISADVANTAGED STUDENTS IN GRADES 3, 5, AND 8

ABSTRACT

The purpose of this causal-comparative study was to examine the impact of selected research-based whole-school reading reform programs, Success for All (SFA), Core Knowledge, and Direct Instruction (DI) on the Criterion Referenced Competency Test (CRCT) reading scores of African-American students and students from disadvantaged subgroups in Grades 3, 5, and 8 in a large urban school system. The target sample consisted of 61 elementary schools and 16 middle schools. Student data consisted of 3,533 data points for African-American students and 7,550 data points for disadvantaged students attending reform and non reform schools respectively. A causal-comparative research design was the methodology employed. Chi-square and Mann-Whitney statistical techniques were used to test 13 hypotheses to determine if there were any significant differences between the reading proficiency of schools and students with reforms compared to those without reforms using race (African American) and socioeconomic status (Disadvantaged) as control variables. The results did not provide any evidence of differences in reading proficiency between reform and non-reform at the school level. However, reading proficiency differed significantly at the student level where Disadvantaged students attending schools with reading reforms showed significant improvements. Reading proficiency rates differed significantly between reform models. Significant proficiency rates were found in schools implementing Direct Instruction or Core Knowledge when compared with the Success for All model.

Acknowledgments

First of all, I am grateful to the Almighty God for equipping me to complete this dissertation and providing me with a family who literally carried me through this process. I like to say thank you to my wife, Myrja who supports me in all of my endeavors and has done a wonderful job carrying out the vision we have for our home and children, Erica and Macy. Secondly, I would like to thank my parents Drs. Chuck and Jean Fuller who continue to model the way and provide unending guidance and wisdom. My gratitude towards my parents throughout the dissertation process cannot be expressed in words alone. Thirdly, I would like to acknowledge Mrs. Myrtice C. Johnson, my mother-in-law for undergirding our efforts by picking up our daughters from school and providing childcare whenever we asked thereby allowing me to write and study without concern for their well being.

I would like to express the deepest appreciation to my committee chair, Dr. Rollen Fowler who has been accessible and constructive throughout the entire dissertation process. His insight and expertise on educational reforms has been invaluable and appreciated.

I would like to thank my committee members, Dr. Amy McLemore and Dr. Chanelle Sweet-White, whose work demonstrated to me that concern for others should never be snubbed out by the hectic nature of academia and family.

Finally, I would like to thank Gladis Richardson (deceased), my 8th grade math tutor who spent several Saturdays with me revealing the beauty of Algebra. I still have a vivid picture of Gladis rolling pencils in her hands as she waited patiently calling me her “beloved” while I worked through a multitude of math problems. The guidance and love of many has truly blessed me exponentially.

Table of Contents

ABSTRACT.....	iii
Acknowledgments.....	iv
List of Tables.....	viii
List of Abbreviations.....	ix
CHAPTER ONE: INTRODUCTION.....	1
Background.....	1
Problem Statement	5
Purpose Statement	6
Significance of the Study.....	9
Research Questions and Null Hypotheses.....	10
Identification of Variables.....	14
Success for All.....	14
Direct Instruction.....	14
Core Knowledge.....	15
Definitions.....	15
CHAPTER TWO: REVIEW OF THE LITERATURE.....	17
Introduction.....	17
Theoretical and Conceptual Framework.....	17
Historical Perspective.....	19
Related Literature.....	21
Poverty and Achievement.....	21
Achievement Gaps.....	24
Whole-School Reform Programs.....	29
Research-Based Reforms.....	33

African-American Achievement Gap Contributing Factors.....	35
Success for All.....	38
Direct Instruction.....	41
Core Knowledge.....	44
Variables Impacting Reform Success/Failure.....	45
Summary.....	46
CHAPTER THREE: METHODOLOGY.....	49
Design.....	50
Threats to internal and external validity.....	50
Research Questions and Null Hypotheses.....	51
Participants.....	54
Setting.....	57
Instrumentation.....	58
Validity.....	59
Reliability.....	61
Procedures.....	63
Data Analysis.....	63
CHAPTER FOUR: FINDINGS.....	67
Analysis Procedures.....	67
Research Questions.....	68
School-Level Analysis.....	70
School-Level Results Grade 3.....	71
School-Level Results Grade 5.....	72
School-Level Results Grade 8.....	74
Student-Level Analysis Findings Grades 3, 5, 8.....	76

Student-Level Findings Grade 3.....	76
Student-Level Findings Grade 5.....	78
Student-Level Findings Grade 8.....	80
Comparison of Reform Models.....	81
Whole-School Comparison Findings.....	81
Direct Instruction versus Success for All.....	82
Core Knowledge versus Success for All.....	82
Core Knowledge versus Direct Instruction.....	83
Summary.....	84
CHAPTER FIVE: SUMMARY AND DISCUSSION.....	86
Summary of Findings and Discussion.....	87
Summary and Discussion of School Level RQ 1, 2, 3 and Ho 1-6.....	87
Student-Level Summary and Discussion of RQ 4, 5, 6 and Ho 7-12.....	91
Connection to the Literature.....	93
Summary and Discussion of RQ 7.....	94
Summary and Discussion of the comparison of Reform Models.....	95
Limitations of the Study.....	96
Implications.....	99
Recommendations for Future Research.....	100
Conclusions.....	101
REFERENCES.....	104
APPENDIX A: IRB APPROVAL.....	130
APPENDIX B: TABLES ON NUMBER OF STUDENTS TESTED.....	131

List of Tables

Table 1: District and School Level Demographics	58
Table 2: Reliability Coefficients (Cronbach’s Alpha) and Raw Scores SEM for Subject/Tests by Grade.....	63
Table 3: Grade 3 Proficiency Rates for Reform and Non-Reform Schools (School Level Analysis).....	73
Table 4: Grade 5 Proficiency Rates for Reform and Non-Reform Schools (School-Level Analysis).....	75
Table 5: Grade 8 Proficiency Rates for Reform and Non-Reform Schools (School-Level Analysis).....	77
Table 6: Grade 3 Proficiency Rates for Reform and Non-Reform Schools (Student-Level Analysis).....	79
Table 7: Grade 5 Proficiency Rates for Reform and Non-Reform Schools (Student-Level Analysis).....	80
Table 8: Grade 8 Proficiency Rates for Reform and Non-Reform Schools (Student-Level Analysis).....	82
Table 9: Proficiency Rates for (DI) Schools and (SFA) Schools.....	83
Table 10: Proficiency Rates for Core Knowledge (CK) and Success for All (SFA) Schools.....	84
Table 11: Proficiency Rates for Core Knowledge (CK) and Direct Instruction (DI) Schools.....	85
Table 12: Sample Population of African-American and Disadvantaged Students in Grades 3, 5 and 8.....	89

List of Abbreviations

AMO.....	Annual Measurable Objective
AYP	Adequate Yearly Progress
CK	Core Knowledge
CRCT.....	Criterion Referenced Competency Test
DI.....	Direct Instruction
NCLB	No Child Left Behind
SFA.....	Success for All
SPSS	Statistical Package for the Social Sciences

CHAPTER ONE: INTRODUCTION

The No Child Left Behind Act (NCLB; 2001) is a national initiative designed to improve the overall academic achievement of all United States students by requiring states to demonstrate continuous improvement. The legislation is currently being considered for reauthorization by the President and the Congress (Present, 2010). NCLB (2001) contains specific mandates detailing academic subjects, grade levels, and proficiency standards that states must implement, as well as consequences associated with not fulfilling the aforementioned requirements.

Although the impact of NCLB (2001) has been somewhat mixed (Baker, Rieg, & Clendaniel, 2006; Gersten et al., 2009), the majority of the research appears to suggest that the initiative has not and is not achieving its intended results to the extent expected, especially as it relates to narrowing the achievement gaps between minorities, the economically disadvantaged and students with disabilities (Porter, 2007). There are also concerns regarding how local school districts and schools select and implement research-based reform initiatives to improve the academic performance of schools not meeting the adequate yearly progress (AYP) requirement stipulated by NCLB (2001). In many instances, school districts and individual schools continue to use reform programs that have not proven to improve or close the achievement gap between minorities, disadvantaged, or students with disabilities (Forsee, 2010).

Background

The NCLB Act was signed into law by President George W. Bush in January of 2002 (Anyon & Greene, 2007). The goal of NCLB (2001) is to have all students proficient in language arts and math by 2013-2014. By most measuring devices, NCLB (2001) is one of the most comprehensive U.S. education policies over the past 50 years relative to the influence of the Federal Government and national educational funding policies (Rand, 2007). The mandate of

NCLB encompassed all states and all public schools and therefore significantly altered how schools operated from 2002 to the present (Borman et al., 2010). One of the central mandated components of the legislation is the establishment of academic benchmarks and annual assessments of student progress.

Unlike NCLB's predecessor, Goals 2000, where the focus was on national standards, outcome-based education, and support and training for teachers, the NCLB framework is significantly more focused on accountability and consequences for failure to meet the required annual achievement benchmarks and objectives. NCLB accountability initiatives consist of Adequate Yearly Progress (AYP) and Annual Measurable Objectives (AMO). Specifically, NCLB (2001) requirements include (a) annual assessments, (b) AYP, (c) AMO, (d) participation, (e) subgroups, (f) graduation rates, and (g) average daily attendance.

According to the U.S. Department of Education (2010), NCLB contains four guiding principles:

- Schools are expected to teach students using methods proven to be successful.
- Schools and districts must demonstrate that all students are making academic progress.
- Teachers and paraprofessionals (such as classroom aides) must be highly qualified to work with students.
- Parents are given more information and more choices concerning their child's education.

NCLB's first guiding principle—that schools will utilize proven, research-based instruction to teach students—will guide this proposed research. NCLB (2001) specifically requires schools to utilize proven teaching methods and programs to ensure all students are performing at or above the expected level of proficiency. The act further requires schools to

utilize scientifically-based research programs and teaching methods to assist in closing the achievement gap for minorities and students from disadvantaged subgroups. NCLB (2001) regulations do not stipulate which research-based programs or teaching methods must be used. In effect, each state is left to select any research-based or scientifically-based reform program deemed appropriate. As a result of the lack of federal standards or criteria, states, local school districts, and individual schools have different programs, making objective comparisons relative to the impact of these programs on academic achievement difficult, if not impossible (Armstrong, 2010; Harford, 2007; King, 2011).

Although each of the aforementioned requirements are important, one of the primary goals of NCLB is to narrow class and racial achievement gaps among Whites, African Americans, and other ethnicities. The research is clear in concluding there has been and continues to be a significant correlation between academic proficiency, race, and socioeconomic status (Orr, 2003). In short, on a percentage basis, there are proportionately more disadvantaged minority students than White disadvantaged students (Williams, 2011). In effect, the NCLB directs particular attention to requiring local school districts to place a great deal of emphasis on closing the persistent achievement gaps of traditionally underserved subgroups, such as specific minorities and low socioeconomic groups (Anyon & Greene, 2007; Berliner, 2006). Both the theoretical and conceptual frameworks for this study have historical and current underpinnings regarding how education is best delivered. Historically, minorities and other subgroups, such as the disadvantaged and students with disabilities, have been underserved, resulting in significant disparities in academic achievement (Harford, 2007).

The framers of a Nation at Risk, Title I, and NCLB were painfully aware of these disparities when they created the aforementioned educational reform legislation (Paris, 1994;

U.S. Department of Education, 2007). As such, the framers paid particular attention to including special provisions for providing components to each of the respective pieces of legislation in an attempt to remedy or eliminate the aforementioned disparities. Although each of the previously mentioned policies helped to narrow the achievement gaps between the majority population, minorities, and the disadvantaged, achievement disparities for minorities and the disadvantaged persist (Berliner, 2006; Chapman, 2007).

For example, studies indicate that, in general, African-American students attend the lowest performing schools in most states. Further, African Americans and other minorities consistently score lower on most indicators of academic success than White students (Archibald, 2006). Moreover, 72% of African Americans attend high-poverty schools. Lastly, the percentage of African Americans meeting proficiency in national assessments in reading and math is less than one fourth of White students (Kin & Sunderman, 2005).

The impact of NCLB (2001) on African-American students and disadvantaged students has been mixed. The data provided by the National Assessment of Educational Progress (2011) and The National Center for Fair and Open Testing (2011) has suggested that the disparities between increased state performance scores and decreased NAEP performance since NCLB indicate that NCLB has failed to improve overall performance, especially in the subject areas of reading and mathematics. Equally important is the finding of the Trial Urban District Assessment that showed little evidence of improvement in math and reading since NCLB became law (Samuels, 2009; Yudof, 2008). In addition, all of the sources cited indicated that only modest closing of the gaps between racial minorities and Whites or between disadvantaged and non-disadvantaged students has occurred (American Psychological Association, 2011). Although several states have experienced a narrowing of the achievement gaps among African-

American students, the majority of the evidence indicates that the gains have been positive but small (Center on Educational Policy, 2012). The evidence further suggested that the overall disaggregated data continue to support the contention that racial minorities and other underserved subgroups are not making academic progress on par with the majority population and non-disadvantaged students (Taylor, 2005).

Problem Statement

The problem statement is that NCLB (2001) requires each school to ensure every student demonstrates measurable achievement gains annually, specifically in reading and mathematics that support academic progress. The act further mandates school districts pay particular attention to narrowing the achievement gap between subgroups that have traditionally been underserved. Research has shown that African-American and disadvantaged students often attend the lowest-performing schools and score significantly below White students on almost every achievement indicator (Braun, Chapman, & Vezzu, 2010). Moreover, NCLB stipulates that school districts must utilize scientifically-based reform initiatives and rigorous teaching methods to address the academic needs of African-American students and students from disadvantaged populations. In addition, districts must provide annual assessments of these programs relative to the program's effectiveness for improving achievement across the total school population, but specifically demonstrate positive achievement impact to the aforementioned subgroups of African-American and disadvantaged students. More often, school districts employ expensive reform initiatives but seldom perform either formative or summative program evaluations to justify their existence (Mandinach et al., 2005; Myers-Miller, 2010). In short, the scientifically-based programs must generate results that are valid and replicable (Duncan & Magnuson, 2005; Hess, 1999; 2004). In view of the fact that the current NCLB is under consideration for reauthorization and the urban

school system addressed in the research study currently has six educational reform models operating, it is imperative that the impact of each of the respective reform models be evaluated relative to the model's impact on student achievement. It is also important that an evaluation of each of the reform programs be made relative to which program was more effective in producing the greatest achievement improvement. There is currently little formative or summative program evaluation data that provides a base for an objective quantitative assessment of program effectiveness or a relative comparison of each reform program.

Purpose Statement

This causal-comparative research study has two purposes. The primary purpose was to examine the impact of selected research-based, whole-school instructional reading reform programs on the academic proficiency of African-American students and students from disadvantaged subgroups in reading in Grades 3, 5, and 8. Academic proficiency, for purposes of this study, is defined by the combined percentages of students who perform at the "meets" and "exceeds" categories on the reading subsection of the Criterion-Referenced Competency Test (CRCT). The use of aggregated student data is both appropriate and efficient where the intent is to produce a whole-school or comprehensive assessment as opposed to individual or diagnostic assessments. In addition, NCLB requires each state to collect and use aggregated data to meet accountability requirements. More specifically, the research attempted to determine if and to what degree the respective reform programs significantly improved the CRCT test scores of African-American and disadvantaged students compared to students in schools where the selected whole-school reforms were not implemented. In effect, the researcher examined reading proficiency for 2010-2011 across schools for the purpose of determining if any significant performance changes occurred based upon the respective whole-school reform implemented.

The second purpose of this study was to compare the relative impact of the aforementioned selected instructional whole-school reform programs for the purpose of determining which of the three whole-school reading reform programs produced the greatest percentage of students scoring in the meeting or exceeding reading levels improvement on the CRCT for African-American and disadvantaged students in Grades 3, 5, and 8 in a large urban school system. The selected initiatives are the following educational reforms:

Success for All (Independent Variable)

Success for All Foundation is founded on the belief that every child can and will learn. No matter what the academic level of the student, each child is challenged to do his or her best and the contributions of all team members are equally valued. Cooperative learning is one of the most powerful tools teachers have in providing the level of engagement and academic and social support their students need to be successful. In the cooperative learning classroom, all students benefit from the constant coaching, encouragement, and feedback of their peers. More of the responsibility for learning rests on students and teams; therefore, teachers are able to spend more time working with individuals and small groups of learners, doing the kind of teaching that originally drew them to the field (Borman et al., 2006; Center for Research and Reform in Education, 2005).

Direct Instruction (Independent Variable)

Direct Instruction (DI) includes explicit, systematic instruction based on scripted lesson plans. Students are grouped and re-grouped based on their rate of progress through the program. Direct Instruction emphasis is based on the efficiency of instruction. DI programs are meant to accelerate student progress; therefore, lessons are designed to bring students to mastery as quickly as possible. Frequent curriculum-based assessments

help place students in ability groups and identify students who require additional intervention as well as embedded professional development/coaching. DI programs may be implemented as stand-alone interventions or as part of a school-wide reform effort. (Lindsey, 2010, p 2)

Core Knowledge (Independent Variable)

The Core Knowledge Language Arts Program (Bradley, 2005) combines systematic phonics-based instruction in decoding skills with extensive read-aloud to build both oral language and background knowledge-word knowledge and world knowledge. Although each of the aforementioned reform programs has different strategies, activities, and processes, all are designed to provide research-based instructional strategies to improve reading proficiency. In fact, it is the differences between the respective reforms that provide both content and internal validity to their impact on academic improvement.

This research study makes the assumptions that each of the respective reform models were implemented for a minimum of 3 years and all staff members received the necessary training relative to the reform model purpose and instructional strategies. Another assumption is all students classified as African American meet both the state and local school district's criteria for the classification. The same assumption is held for students classified as being disadvantaged. Finally, the term *disadvantaged* maybe based upon several qualifying criteria. These criteria may include income threshold, number of persons in a family, and educational level of parents. For purposes of this study, disadvantaged students will be based on the student's eligibility to receive free or reduced priced meals (as reported to the Georgia Department of Education [GA DOE] in October 2008 Nutrition count).

Although each of the respective educational reform programs is designed to impact reading performance, it is understood that all instruction is impacted by the student's reading ability. The researcher also recognizes the limitation of not being able to ensure that all teaching staff possessed the same level of experience or prior subject area knowledge and skills. As a result, the researcher risks not being absolutely certain that each of the respective reform programs will be delivered at the same level and could impact the student's acquisition of the materials.

Significance of the Study

The significance of this study is based upon several historical and contemporary factors. Foremost among the current factors is the fact that the NCLB (2001) is currently being considered for reauthorization. Major changes and/or revisions are being proposed and considered (Harford, 2007). Issues such as the use of national standardized assessments, reductions in both scope and funding formulas may have a profound impact on the future of the program. Giving states and individual school districts greater control over subject areas to focus on and which subgroups to provide additional resources to are issues that may challenge state and local school districts (Duncan & Magnuson, 2005).

The results of this study will provide states and school districts with additional reform programs that have been tested and validated under field conditions. In addition, this research study specifically evaluates reform programs designed to improve the academic performance of African-American and disadvantaged students. Further, the results will help inform the existing literature relative to programs that have been shown to be effective in closing the achievement gap for the disadvantaged subgroups. Finally, the results of the study will provide schools and other stakeholders with much needed data-driven formative and summative data relative to the degree to which each respective reform program can impact achievement in reading.

Research Questions and Null Hypotheses

Research Question 1: Is there a significant difference in the school level third grade Reading performance between African American or disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African- American or disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho1: There is no significant difference in the school level third grade Reading performance between African American students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African American students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

Ho2. There is no significant difference in the school level third grade Reading performance between disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

Research Question 2. Is there a significant difference in the school level fifth grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho3. There is no significant difference in the school level fifth grade reading performance between African-American students in the intervention group and African-American students in the comparison group.

Ho4. There is no significant difference in the school level fifth grade Reading performance between disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

Research Question 3. Is there a significant difference in the school level eighth grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho5: There is no significant difference in the school level eight grade Reading performance between African American students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African American students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented

Ho6: There is no significant difference in the school level eight grade Reading performance between disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

Research Question 4

Is there a significant difference in the student level third grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American or disadvantaged students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho7. There is no significant difference in the student level third grade reading performance between African-American students in the intervention group and African-American students in the comparison group.

Ho8: There is no significant difference in the student level third grade reading performance between disadvantaged students in the intervention group and disadvantaged students in the comparison group.

Research Question 5. Is there a significant difference in the student level fifth grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and disadvantaged students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho9. There is no significant difference in the student level fifth grade reading performance between African-American students in the intervention group and African-American students in the comparison group.

Ho10. There is no significant difference in the student level fifth grade reading performance between disadvantaged students in the intervention group and African-American students in the comparison group.

Research Question 6. Is there a significant difference in the student level eight grade reading performance between African-American students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented

Ho11. There is no significant difference in the student level eighth grade reading performance between African-American students in the intervention group and African-American students in the comparison group.

Ho12. There is no significant difference in the student level eighth grade reading performance between disadvantaged students in the intervention group and disadvantaged students in the comparison group.

Research Question 7

Is there a significant difference between the relative impact of Success for All, Direct Instruction, or Core knowledge on the reading proficiency of students following a pairwise comparison between each of the reform models?

Ho13. There will be no significant difference between the relative impact of Success for All, Direct Instruction, or Core Knowledge on the reading proficiency of students following a pairwise comparison of each of the reform models.

Identification of Variables

The independent or treatment variables for this study were three whole-school reform programs listed below:

Success for All (independent/treatment variable). The Success for All Foundation is founded on the belief that every child can and will learn. No matter what the academic level of the student, each child is challenged to do his or her best, and the contributions of all team members are equally valued. Cooperative learning is one of the most powerful tools teachers have in providing the level of engagement and academic and social support their students need to be successful. In the cooperative learning classroom, all students benefit from the constant coaching, encouragement, and feedback of their peers, since more of the responsibility for learning rests on students and teams, teachers are able to spend more time working with individuals and small groups of learners, doing the kind of teaching that originally drew them to the field (Borman et al., 2006; Center for Research and Reform in Education, 2005).

Direct Instruction (Independent/treatment variable). Direct Instruction (DI) includes explicit, systematic instruction based on scripted lesson plans. Students are grouped and re-grouped based on their rate of progress through the program. Emphasis is based on efficiency of instruction. DI programs are meant to accelerate student progress; therefore, lessons are designed to bring students to mastery as quickly as possible. Frequent assessment and curriculum-based assessments help place students in ability groups and identify students who require additional intervention. Embedded professional development/coaching. DI programs may be implemented as stand-alone interventions or as part of a school-wide reform effort (Lindsey, 2010, p. 25).

Core Knowledge (Independent/treatment variable). The Core Knowledge Language Arts Program combines systematic phonics-based instruction in decoding skills with extensive read-aloud activities to build both oral language and background knowledge-word knowledge and world knowledge. The Core Knowledge curriculum is based upon the theory that learning is the basis for learning and children are exposed to specific content within each respective subject area.

Definitions

Adequate yearly progress (AYP): This is the minimum level of proficiency that school districts and schools must achieve each year as determined by the No Child Left Behind Act. To make AYP, a school and district must meet the required participation rate, the AMO, in language arts and math (Georgia Department of Education, 2010).

Annual measurable objective (AMO): This is the percent of students who must be proficient on the above exams as required by the state (Georgia Department of Education, 2010).

African-American students: A person having origins in any of the black racial groups of Africa and not of Hispanic origin (U.S. Census Bureau, 2011).

Annual assessments: The annual achievement test(s) required by NCLB. Each state determines its own standardized assessment instrument.

Disadvantaged student: Calculated by dividing the number of students eligible to receive free or reduced-priced meals (as reported to the Georgia Department of Education in October 2008 nutrition count) by the total school enrollment (as reported by October 2008 FTE count). The Georgia Department of Education classifies each student's socioeconomic status by determining whether or not the respective student is eligible for free or reduced school lunch. Achievement scores are provided for each respective socioeconomic status category.

Georgia Criterion Referenced Competency Test (GCRCT): The Georgia Criterion Referenced Competency Tests (2011), is a set of tests administered at public schools in the state of Georgia designed to test the knowledge of first through eighth graders in reading, English/language arts (ELA), and mathematics; third through eighth graders are also tested in science and social studies.

No Child Left Behind: NCLB (2001) supports standards-based education reform, which is based on the belief that setting high standards and establishing measurable goals can improve individual outcomes in education. NCLB contains specific components requiring subject area, grade-specific assessments in order for states to qualify for Federal reimbursements (U.S. Department of Education, 2007a, 2007b).

Proficiency level: The three levels used to measure student performance on the CRCT are defined as: Three (3) Exceeds Expectations, Two (2) Meets Expectations, and One (1) Does Not Meet Expectations (Georgia Department of Education, 2011).

Subgroups: The subgroups that must meet the above standards, as currently defined by the State of Georgia, are Caucasian, African American, Asian, Native Alaskan, American Indian, Hispanic, Economically Disadvantaged, Students with Disabilities, and English Language Learners.

Whole-school reform: This is a systematic approach towards improving achievement by taking into account the entire primary factors existing in a typical school. The factors include curriculum, manpower, organizational structure, and policy matters (Rosenshine, 1996).

CHAPTER TWO: REVIEW OF THE LITERATURE

Introduction

Review of the literature related to NCLB (2001) supports the belief that improving the proficiency of African-American students and students from disadvantaged populations through selected reform intervention programs is essential to improving schools at all levels (Braun, Chapman, & Vezzu, 2010; Duncan & Magnuson, 2005; Pogrow, 2002). The review of the literature on selected reform programs is organized into seven sections and subsections. Section one provides an overview of current legislation. Section two explores the theoretical framework of the research problem as evidenced by the literature. Section three contains a historical perspective of the issues that provided the framework of how and why narrowing the achievement gaps for African-American and disadvantaged students is an integral component of NCLB. Section four provides a critical review and analysis of the research surrounding efforts (reform initiatives) to narrow the achievement disparities of African-American and disadvantaged students. Section five summarizes research findings and factors that contribute to the literature surrounding reading proficiency and the success or failure of reform initiative (Archibald, 2006; Clewell, Campbell, & Perlman, 2007; Pogrow, 2002). The sixth section of the literature review provides an in-depth description and analysis of each of the research-based reading reform programs (e.g., Success for All, Direct Instruction, and Core Knowledge) that serve as the independent variables of the study. The final section of the literature review provides a summary and critical review of the related literature.

Theoretical and Conceptual Framework

The theoretical framework for this research focuses upon two concepts that are central to understanding the purpose of the study. The research addresses the potential impact of research-based educational initiatives on improving the academic proficiency of African-American

students and students from disadvantaged populations. Most of the so-called theories regarding race and academic performance are not theories at all, but are loosely associated with social or cultural theories. For example, the thesis of Woodson (2009) was that “African Americans of his day were being culturally indoctrinated, rather than taught, in American schools” (p.25).

Woodson (2009) further theorized that African Americans are disassociated from a quality education due in part to not being taught their history and, as a result, they have difficulty understanding many of the cognitive aspects of learning.

Other theories surrounding race and achievement assume a more insidious explanation of the achievement gaps between African Americans and Whites. Researchers such as Herrnstein and Murray (1994) suggested that academic achievement is greatly influenced by genetics and is therefore mostly resistant to any significant improvement. This explanation, however, opposes of theories advanced by the majority of researchers and educators who adhere to social-cultural theoretical explanations Seaton (2010). According to Sadovnik, Cookson, and Semel (2001), “The contention that race is related to educational outcomes are undeniable, although given the nature of U.S. society, it is extremely difficult to separate race from class” (p. 1). Sadovnik (2010) further suggested that the achievement gap between minorities and Whites is not complex, given that minorities do not receive the same educational opportunities as Whites and their rewards for educational attainment are significantly less.

The theoretical framework regarding the definition of poverty or being disadvantaged is both dubious and complex. The literature is replete with various definitions and categories of criteria for what constitutes being disadvantaged (Hagenaars & Klass, 1988; Cutler, 1984; de Vos, 1988). According to Hagenaars and Klass (1988), “The choice of a specific poverty definition as the one and only measuring rod thus appears to have major consequences, both for

the observed incidence of poverty and for the distribution of the poor over social subgroup.” The implications expressed by some educators, theorists, or people in general, is that being disadvantaged or poor has negative connotations that cut across social and cultural boundaries (Bowen & Bok, 1998; West, 2001) Low academic performance is one of the most prominent associated conditions typically examined relative to the disadvantaged.

There are several theories associated with poverty or disadvantage. Only three, for purposes of this study, have relevance. The culture theory of poverty suggests that individuals living in poverty have been psychologically impacted by their environment and are therefore not likely to realize that there are other alternatives and other environments (O’Connor, 2001). The causal theory of poverty postulates that disadvantaged persons cause their own poverty due to a host of factors, including low self-esteem and a lack of motivation. Herrnstein and Murray (1994) describe several aspects of their conceptual and theoretical views of poverty. Essentially, the authors provide what they consider research-based evidence that being disadvantaged is highly correlated with the level of intelligence and as a result, poor people will never be successful academically due to both genetic and environmental factors. Much of the research presented in the following sections provides an abundance of research that contradicts Herrnstein and Murray’s conclusions (Seaton, 2010; Okayed, 2009; Monastersky, 2008; Webber, 2001).

Historical Perspective

Initially, the national movement to take a serious look at the state of education in the country was not caused by a concern for the disparities between the academic achievement of minorities or other subgroups. The initial call to arms was sounded by the *Nation at Risk* report (1983) that did not focus specifically on the education of minorities, the disadvantaged, or students with disabilities (Bensalem, 2011). The report provided a clear picture of the poor state

of affairs within elementary and secondary education, ranging from a general lack of basic reading comprehension rates and high dropout rates (Encyclopedia of American Education, 2011). Still, the far-reaching recommendations resulting from the *Nation at Risk* Report (U.S. Department of Education, 2010) marked the beginning of what was referred to as the first wave of educational reforms. Further, there was no mention or details of funding whole-school reforms that are currently provided by NCLB (2001).

Among the reforms included in the report's recommendations was to establish content standards for higher education. The standards included four years of English, three years of mathematics, science, social studies, and one-half year of computer science (Thattai, 2008). The commission also recommended local school districts should consider longer school days and school year (U.S. Department of Education, 1983). Again, there was no reference to addressing the achievement gaps of minorities, the disadvantaged, or students with disabilities.

The major legislation, however, that resulted in a major shift in the improvement of American education was the passage of the Improving America's Schools Act in 1994 (McLaughlin, 1995). This act attempted to establish a model for identifying proven systemic remedies to gaps in student achievement, especially for minorities and other underserved groups of students. Nevertheless, this attempt fell short of correcting the achievement gaps for several reasons. Individual schools and entire school districts struggled to find instructional approaches and programs that would directly impact student achievement on a large scale as mandated by the new federal regulations that existed under Comprehensive School Reform (Encyclopedia of American Education, 2011; Kim & Sunderman, 2005).

The *Goals 2000: Educate America Act* was enacted approximately 10 years following the initiatives implemented as a result of the *Nation at Risk* report (Cohen, 2003). Similar to the

recommendations from the *Nation at Risk* report, the *Goals 2000* Act was based on the belief that higher standards and expectations would result in higher achievement. The legislation did not contain strong accountability components and did not include consequences for noncompliance or failure to reach specified goals (Paris, 1994). However, the legislation recommended systemic reforms for education of state and local districts (College Board, 2007). The legislation did not provide specifics regarding which reforms would be appropriate or how their effectiveness would be measured. Nevertheless, there were no mandated recommendations regarding the achievement gap between minorities, the disadvantaged, and White students who were not disadvantaged. Approximately another 10 years passed before *No Child Left Behind* (2001) was signed into law containing specific initiatives requiring schools to pay particular attention to the disparities in achievement for students who were traditionally underserved, such as minorities, low-income students, and students with disabilities (Clewell et al., 2007; Dee & Jacob, 2010).

Related Literature

Poverty and Achievement

The understanding that forms the structure underpinning NCLB (2001) is the requirement that providing a quality education for all students, especially students from minority and economically disadvantaged groups, is essential to achieving the core purposes of the Act. The theoretical framework is, therefore, derived and driven by the long standing belief that all children can learn if provided with the necessary opportunities and resources (Carlyle, 2008). NCLB (2001) changes the federal government's role in kindergarten through Grade 12 education by focusing on school success as measured by student achievement. The NCLB legislation of 2001 (U.S. Department of Education, n.d.) contains four basic education reform principles that

include (a) Stronger accountability for results, (b) Increased flexibility and local control (c) Expanded options for parents; and (d), An emphasis on teaching methods that have been proven to work.

More specifically, a stronger accountability for results and an emphasis on teaching methods that have proven to work is the theoretical foundation for the current study that focus upon the impact of research-driven educational reform programs that have been proven to work. Further, it has been well established through research that certain subgroups have been subjected to conditions that were inherently unequal relative to access to quality educational opportunities (Avila, 2010; Reardon, 2011). A host of research studies have indicated ethnicity and socioeconomic status have had a major impact on the academic achievement of many students in the subgroups mentioned above (Clewell et al., 2007; Rist, 1973). For example, there can be no denial that past cultural and social perception by some that students from minority communities are somehow less able intellectually to handle the rigor required performing at high academic levels in areas of reading and mathematics (Kulm, 2007). A large body of research has addressed the interaction and correlation between poverty, ethnicity, and academic achievement (Strayhorn, 2010; Yeung & Conley, 2008). The essence of much of the literature is divided between viewing poverty as a condition and viewing poverty as a cause of academic failure. This philosophical divide is based in part upon the persistent achievement gap between various indicators of poverty, ethnic and racial groups (Charles, Fischer, Mooney, & Massey, 2009).

Poverty, as a construct, is generally defined as the absence of the basic needs required to afford individuals or groups basic human necessities such as food, clothing, and shelter (U.S. Census Bureau, 2009). This type of poverty is generally referred to as *absolute poverty*. Another type of poverty, *relative poverty*, is defined as a condition where individuals or groups have

fewer resources, on average, than the greater population. Relative poverty, typically referred to as low SES, is generally based upon the relative economic wealth of an individual or family. In short, it is a combination of income and the number of family members in a household, coupled with the age of the householders that results in the creation of a poverty threshold (Armstrong, 2010; Baker, 2009; U.S. Census, 2009). About 1.7 billion people worldwide live in absolute poverty. Relative poverty, poverty based primarily upon income level, best describes the major type of poverty in the United States (U.S. Census, 2009). The number of people living in relative poverty in the United States increased by 42.9% over the past 10-year period. Approximately 14.3% of the U.S. population had income below the respective poverty thresholds in 2009 (U.S. Census Bureau, 2009).

Several research studies within the past five years clearly have indicated that poverty, among all identified variables, is the most impactful factor on the academic achievement of students (Burney & Beilke, 2008; Duncan & Magnuson, 2005; Kim & Sunderman, 2005; Levin, 2007; Monastersky, 2008). Other research studies have been more specific in suggesting that poverty is a major contributing factor to the gaps in educational achievement; however, when coupled with minority group affiliation, the gap becomes significantly wider (Braun et al., 2010; Hughes, Manns, & Ford, 2009). Eddy and Easton-Brooks (2011) found that students from minority populations, specifically African Americans and Hispanics, are disproportionately represented in the bottom percentile of all of the core academic subjects, such as reading, mathematics, and science (Robelen, 2009; Tan & Barton, 2010).

The relationship between SES and achievement is further validated by the most recent findings from the U.S. Department of Education (2008) and the National Center for Education Statistics (2011). According to the data provided by the National School Lunch Program

(NSLP), the percent of students eligible for free or reduced-price school lunch increased from 33% in 2003 to 43% in 2010 (U.S. Department of Education, 2010). More importantly, the same report indicated there is a strong relationship between the academic achievement of students in reading and mathematics based upon whether students were eligible for free or reduced-price school lunch. Again, according to NAEP (2011), students not eligible for free or reduced-price school lunch (not disadvantaged) had an average reading score of 235 compared to a score of 206 for disadvantaged students eligible for free school lunch (U.S. Department of Education, 2010).

Achievement Gaps

Although the research pertaining to NCLB (2001) and the impact on minorities and the disadvantaged remains mixed, the majority of recent data concludes that the achievement gap, especially between African-American and White students, has not changed significantly since the enactment of NCLB (Evans, 2005; 2005; Freeman, 2005; Orr, 2003; Yeung & Conley, 2008). The statistical evidence is striking. For example, according to data released by the U.S. Department of Education (2010), the majority of African-American students attend the lowest performing schools and score lower on standardized tests than their counterparts. Further, the demographic data indicates that 72% of African-American students attend high-poverty schools and the percentage of African Americans meeting proficiency on national assessments in reading and math is less than one fourth of White students (Baker, 2005; U.S. Department of Education, 2007).

According to the statistics provided by the National Assessment of Educational Progress (NAEP, 2000), the gap between African Americans and Whites narrowed in both reading and

mathematics. Unfortunately, the improvements were not sustained and the African-American-White achievement gap, once again, begin to widen (Hoff, 2000).

As Sadowski (2001) stated:

Perhaps even more disturbing, these gaps seem to be getting wider each year. Even when researchers control for socioeconomic status, level of parental education, and other factors that contribute to scholastic achievement, the score gap between White and black students persists, and no one is really sure why. (p. 2)

The most recent achievement results, however, reveal another narrowing of achievement gaps between African Americans and Whites since 2009 (NAEP, 2010). The report indicates that the recent improvements occurred in both reading and mathematics; still, the achievement gap between African Americans and Whites remains significant.

Multiple studies have attempted to address the underlying cause of achievement gaps between African Americans and Whites (Burchinal et al., 2011; Desimone, Smith, & Frisvold, 2007)—the belief that early childhood appears to be the starting point for many of the predispositions to future achievement gaps. Burchinal et al. (2011) also added that disparities between African-American and White SES also contribute significantly to the problem. Rothstein (2011) agreed with the assertion that early childhood experiences or lack thereof are factors that must be considered. In addition, Rothstein (2011) indicated that cultural aspects of the early childhood experiences like peer-group associations and other social interactions help to determine the quality of the early childhood experience. Wenglinsky (2004) provided the following explanation for the causes of the achievement gap:

The causes for the achievement gaps are very complicated in nature. According to experts, there are 2 main culprits: The first of which spawns from a student's home life

and includes the student's SES, culture, and background. The second cause is factors related to the schools and teachers. There could be many cultural, home-based, and genetic factors that have had an impact on this discrepancy. (p. 2)

Each of the previously reported causes of the achievement gap provides either directly or indirectly the issue of race, SES, and early childhood experience and culture as major contributors to the achievement gap. Several researchers and educators have suggested that the local school and classroom teachers are also contributing factors (Noguera & Akom, 2000; Wenglinsky, 2004).

The literature examining the impact of being poor or disadvantaged has strongly supported the contention that poverty is the most historically persistent causal factor in academic achievement (Avila, 2010; Carlyle, 2008). In a similar manner, the literature has also addressed the correlation between ethnicity and academic achievement. However, the literature has not done a creditable job of examining or reporting how being both African American and disadvantaged impact academic achievement in spite of the unlimited data that clearly shows that the majority of disadvantaged are African-American people.

Although the research is replete with possible reasons and causes for the difficulties associated with this issue, the problem persists. According to some researchers (Clewell et al., 2007; Harford, 2007; Rothstein, 2008), the causes and reasons are as varied as the students and backgrounds of families and communities. A substantial amount of research has made a connection between being disadvantaged and academic achievement (Braun et al., 2010; Goddard, Salloum, & Berebitsky, 2009; Griffin, Jayakumar, Jones, & Allen, 2010). According to several researchers (Baker, 2009; Hughes et al., 2009; Taylor, 2005), poverty appears not to discriminate based upon race or ethnicity. In short, poverty has a deleterious impact overall.

Although the literature contains studies suggesting a correlation between the impact of being disadvantaged on achievement (Levin, 2007; Okoye, 2009; Rothstein, 2008), there remain several troubling issues that have not been significantly addressed through the literature, thereby giving credence to further exploration of the examination of the impact of being disadvantaged on achievement (Burton, 2007).

Duncan and Magnuson (2005) argued that one cannot read the literature on poverty and academic achievement without, at a minimum, questioning the research approach. In short, much of the literature focuses an inordinate amount of effort studying the issues associated with the impact of being disadvantaged on academic achievement as opposed to addressing specific instructional strategies to limit the negative impact and improve achievement. One research initiative in particular helps to inform the aforementioned theory on the impact of being disadvantaged or in poverty (Ractor, Johnson, & Fagan, 2001). According to Ractor et al. (2001), “The analysis reveals that welfare dependence and single parenthood are the major underlying factors producing child poverty. Race per se is not a significant factor in directly increasing child poverty” (p. 13). Reducing the higher levels of single parenthood and dependence among African Americans is the only way that the level of poverty among African Americans can be reduced (Ractor et al, 2001). In fairness, it is important to note that the research was conducted by the Heritage Foundation (2001) whose self-expressed mission is: “Our mission is to formulate and promote conservative public policies based on the principles of free enterprise, limited government, individual freedom, traditional American values, and a strong national defense” (p. 1). The obvious problem with the conclusions drawn by the report is the fact that it appears cyclical in that it states the single most contributing cause of poverty is single-parent households; yet, they further state that single parent households lead to poverty.

Although the literature suggest that SES is a major contributor to low academic achievement, there is evidence to support the contention that other factors, such as cultural and environmental factors also contribute to the achievement gap between African Americans and Whites (Drukker, Feron, Mengelers, & van Os, 2009). NCLB (2001) appears to have recognized these problems when the legislation was enacted and put forth several mandates designed to address the achievement disparities most pronounced in disadvantaged students (Delpit, 2007; Evans, 2005; Freeman, 2005). It is most unfortunate that the NCLB legislation did not require more oversight and accountability relative to whole-school reforms implemented over the past 10 years of the Act. The fact that the achievement gap for African-American and disadvantaged students persist in spite of the national initiative raises several questions regarding the impact of NCLB (Berlak & Harold, 2001; Kulm, 2007; Williams, 2006).

The literature has also provided data regarding the achievement gap that exists between other ethnicities. For example, multiple studies (Carpenter, Ramirez, & Severn, 2006; Clayton, 2011; U.S. Department of Education, 2010) have provided evidence that the achievement gap between Whites and Asians has been and remains consistent. Nevertheless, some research has shown that the achievement gaps also exist between other ethnicities as well. For example, Birenbaum, Nasser, and Tatsuoka, (2007) examined the achievement gap between Jewish and Arab fourth graders in mathematics. The results indicated that the Jewish fourth graders performed significantly better than did the Arab students. Birenbaum et al. (2011) attributed the differences to cultural and educational resources and also found that gender was also a contributing factor in that boys scored significantly better than girls. Bahr (2010) examined the degree to which remedial mathematics instruction was effective in closing the achievement between Whites, African Americans, and Hispanics. Bahr (2010) found that closing the gap

between White and Asian students was fairly easy compared to closing the achievement gap between African Americans and Hispanics. The differences were due in part to the relatively large achievement gaps between African Americans, Hispanics, and Whites. Bahr (2010) concluded the following:

It is clear that any racial disparities in mathematics preparation and achievement may be attributed to a number of well-documented expressions of socioeconomic inequality, such as academic tracking, lower levels of parental capital, and the poorer quality of primary and secondary schools in neighborhoods characterized by a high percentage of minorities. (p. 212)

Although the literature is replete with research pertaining to the impact of being disadvantaged on African-Americans, Hispanics, and other ethnicities (Bahr, 2010; Orr, 2003; Rothstein, 2008), it is extremely thin with regards to its investigation of how poverty impacts White students. To be clear, the literature has provided a long list of examples on the interaction between race and achievement (Braun et al., 2010; Freeman, 2005); however, few studies, if any, have reported that poor White students consistently outperform middle-class African Americans and other minorities on tests of academic proficiency (Burchinal et al., 2011). Thus, poverty is not only deleterious to African Americans, Hispanics, and other minorities, but apparently has negative effects on Whites as well. This point is significant due to the specific wording in the NCLB (2001), which states special attention must be paid to the underserved subgroup of the disadvantaged student. It does not say directly that the disadvantaged must only be African American.

Whole-School Reform Programs

The concept of whole-school reform, sometimes referred to as comprehensive school reform, came into being initially as the government's response to the 1983 report *A Nation at*

Risk that painted a dismal picture of the state of education in the United States (U.S. Department of Education, 1983). Although there are varying definitions of what constitutes a comprehensive school reform, there is general agreement that comprehensive reforms encompass academics, school management, staffing, and in many instances parent and community involvement (Berends, Bodilly, & Kirby, 2002; Church, 2000; Dunn, 2009).

Actually, the concept of comprehensive school reform existed long before the government officially funded whole-school reform programs. Programs such as Head Start and Follow Through were being implemented as far back as the early 1960s as mechanisms for providing more direct and effective educational services to disadvantaged students (House, Glass, McLean, & Walker, 1978; Lee, Brooks-Gunn, Schnur, & Liaw, 1990; Nehring, 2009). These programs were not considered to be whole-school or comprehensive-reform initiatives because they generally focused on the instruction, curriculum or specific subject areas such as reading and mathematics. More specifically, the general consensus and criticisms of the early reform programs was that they lacked consistency the implemented the programs. Further, according to some, the early programs were often fragmented in the overall approach to addressing the achievement problems of minorities and low-income students (Berends et al., 2002; Williams, 2006).

The implementation of several whole-school reform initiatives that were considered by many to be ineffective lead to the creation of the New American Schools (NAS), a private non-profit organization devoted to assisting school systems improve achievement through the design and implementation of whole-school designs (RAND, 2002; Fischer, 2010; Fullan, 2010). NAS later partnered with RAND for purposes of documenting and analyzing the formative aspects of NAS's efforts. According to RAND (2002), "NAS's purpose was to implement comprehensive,

scientifically research-based whole-school reforms” (para. 2). NAS implemented the program in three phases:

- One-year development phase (1992-1993);
- Two-year demonstration phase (1993-1995, including the 1993-1994 and 1994-1995 school years); and,
- Three-year scale-up phase (1995-1998, from the school year starting in 1995 to the school year ending in 1998). (para. 3)

RAND’s evaluation consisted generally in mixed results. For example, Berends et al. (2002), found the following:

In total, of the 163 schools for which we had data allowing comparisons in performance relative to the district or state, 81 schools (50 percent) made gains in mathematics relative to the district and 76 schools (47 percent) made gains in reading relative to the district. (p. 14)

In a similar effort, the U.S. Department of Education initiated legislation designed to improve the achievement of all students, especially disadvantaged students. The Comprehensive School Reform (CSR) was created and funded to address some of the same issues that NAS and RAND were examining relative to the development of comprehensive, research-based programs. The specific purpose of the legislation was to encourage school systems to implement comprehensive, scientifically research-based whole-school reforms in order to improve student achievement by providing financial incentives (RAND, 2007). The primary requirement was that the selected reforms had to be scientifically research-based instructional practices. One of the overriding problems associated with both the adoption and implementation of many of the comprehensive reform programs was the lack of accountability relative to program evaluation

and assessment (RAND, 2007). Further, the small number of districts that did conduct external program evaluations did not follow established assessment processes. As a result, little merit could be placed on the findings and implications (Feinzimer, 2000). Even in districts where so-called scientifically research-based reform programs were implemented, the programs did not meet validated, evidence-based evaluation processes or procedures. The U.S. Department of Education (2002) attempted to address this deficiency by developing a set of criteria for school district practitioners to use in validating programs purported to be scientifically research-based. The criteria provided by the U.S. Department of Education (2002) are the following: (a) Systematic and empirical, (b) Rigorous data analysis, (c) Reliable and valid data collection, (d) Strong research design, (e) Detailed results that allow for replication, (f) Expert scrutiny.

The U.S. Department of Education (2002), Comprehensive School Reform Program Office, also discusses research designs found to be appropriate for conducting scientifically research-based whole-school reform program evaluations. According to the department's recommendations, all research designs should be either experimental or quasi-experimental designs that utilize random assignment and pre and post measurements where possible.

The research regarding the effectiveness of whole-school reform programs is mixed. One particular example of this mixed-results syndrome is the Learning-Focused Schools (LFS) program. The LSF program specializes in connecting reading comprehension, writing across the curriculum, accelerating and scaffolding learning, balanced literacy, and differentiated assignments with the overall goal of raising achievement (Thompson & Thompson, 2000).

The program was implemented in three elementary schools over a three-year period. The LFS program breaks down instruction into nine reading strategies: (a) main idea, (b) sequencing, compare and contrast, (c) telling how things are alike and different, (d) fact and opinion, (e)

cueing systems, (f) cause and effect, (g) literary elements, (h) inferences, (i) extending thinking activities.

The findings indicated that Hispanic and students from mixed racial groups did not show any statistically-significant improvement in any assessed category using the program. On the other hand, White students and students with disabilities did improve. What was most interesting about the research when exposed to the Learning-Focused Schools Program results is that African-American students performed better than all students overall (Daugherty, 2011; Viadero, 2010).

Another comprehensive educational reform program, AIM at Middle-Grades Results Project, initially funded by the U.S. Department of Education, is a comprehensive school reform program that assists middle-grade schools in becoming high-performing learning and caring organizations (Education Development, 2000). The reform program focuses on the use of “enduring understandings” and “essential questions” to guide student learning that meets local curriculum standards. The middle school, MLK Middle School, was located in a high-poverty district with 90% of the student population eligible for free or reduced-price school lunch. Over the past three years, the student population of 744 students, 90% of whom are eligible for free or reduced-price lunch, have steadily improved their performance on state standardized tests.

Research-Based Reforms

The fact that that NCLB (2001) has not achieved many of the expectations related to closing the achievement gaps among African-American students, disadvantaged students, and students with disabilities has led to a general consensus that the various reforms have not been effective (Fullan, 2010; Grindle, 2004; Neague, 2011). According to Volonino and Zigmond (2007), although many of the reforms may have changed education in general, they may not have

necessarily improved them. They further suggested that some of the reforms may have impacted teaching strategies in a negative way. In contrast, two studies (Present, 2010; Porter, 2007) concluded that although the district made performance gains and made AYP, most of the subgroups—African-American, Hispanic, and disadvantaged students—did not improve significantly. In another study, Avila (2011) concluded that although the English-speaking students did move in a positive direction, the findings indicated that the reform implementation was fragmented and did not improve the academic achievement of minorities or low socioeconomic students. There were similar findings in a study conducted by Fischer (2010) indicating that the accountability system of NCLB (2001) does not explicitly incorporate or hold districts accountable for the skills needed for the twenty-first century (Williams, 2011).

One of the major flaws in the literature is the limited number of comprehensive evaluations of reform initiatives (Adams, 2005; Cruz, 2011; Douglas, 2011; Newman, 2007). Further, the majority of reform evaluations found were conducted by the parent organization, as opposed to an external evaluation contractor. This gap obviously leaves areas of concern to both the validity and reliability of the effectiveness of the program (Delpit, 2007; Educational Research Services, 2001; Fullan, 2010).

There is some criticism regarding the implementation of whole-school educational reforms. Woestehoff (2011) reported that the academic improvements in Chicago were exaggerated and the overemphasis of test results has been a negative initiative. Further, according to Woestehoff (2011), the schools would be better served if the more traditional interventions such as decreased class size and increased parental involvement were implemented.

African-American Achievement Gap Contributing Factors

According to the NAEP Report Card (2011), fourth and eighth grade students improved in reading and mathematics when compared to the 2009 assessments. More specifically, the achievement gap between African Americans and Whites narrowed in reading and mathematics from the previous NAEP assessment in 2009. While the reported narrowing of the gap between African-American and White fourth- and eighth-grade students is a positive sign, it does not negate the fact that the difference between each ethnic group is closer; the fact remains that the relative distance between African-American and White achievement is still statistically significant (U.S. Department of Education, 2010). For example, according to NAEP (2011), the average percentage of White students in the fourth grade scoring proficient in reading was 54% compared to only 15% of African-American students (NAEP, 2011). The previous data indicate that African-American students continue to perform significantly below Whites and even other minorities, such as Hispanic and Asian students.

Although the reasons given for the achievement gap between African Americans and Whites are varied (i.e., poverty, SES, and family structure) (Mickelson, 2009; Rowley & Wright, 2011), the most controversial appears to be racism. Several studies have examined the impact of racism on achievement (Pearlstein, 2009; Purdie-Vaughns, 2004; Sanders, 1997; Webber, 2001). Cohen (2009) conducted a multiple-year study that simply required African-American students to write a reflection on a personal value. Cohen's purpose was to examine how personal self-worth translated into self-affirmation. The simple writing intervention about personal value resulted in improved grade point average and a reduction in the number of times students had to repeat a course (Cohen, 2009). The obvious implication from the aforementioned study is the potential for racism to damage the self-worth of young African Americans.

This phenomenon of racism has been reported in several studies (Pine & Hilliard, 1990; Seaton, 2010; Speight, 2007). Another example of the way in which past racism leaves a psychological “scar” on African Americans is shown in a study conducted by Steele (1997), an African-American psychologist. A standardized test was given to a group of half White and half African-American students. Half of each group was told that the test did not measure ability. The other half was told the test measured academic ability and capacity. The African-American students who were told that the test measured ability scored significantly lower than those who were told the test measured verbal ability. White students scored the same regardless of what they were told about the purpose of the test. The researcher explained the results by suggesting that African-American students were impacted by the long-standing stereotype that African-Americans have limited ability and were therefore negatively impacted. These findings also carry other implications regarding the achievement gap—namely, the impact of racial vulnerability held by African-American students is a part of the subconscious and can be triggered by either verbal or nonverbal factors. The impact of racism is also apparent at the higher education levels. Swim, Hyers, Cohen, Fitzgerald, and Bylsma (2003) and Jones (2011) reported that feelings of hostility exhibited by White students caused African-American students discomfort and increased their level of anxiety relative to academic experiences.

Family structure has also been examined as a mediating variable in the achievement gap between African-American and White students. Cooper, Crosnoe, Suizzo, and Pituch (2010) conducted a longitudinal study to examine the connection between parental involvement, poverty, and children’s math and reading achievement in kindergarten. Cooper et al. (2010) concluded race and socioeconomic factors were second to parental involvement as related to achievement in reading and mathematics. In another study conducted in Montréal, Caldas,

Bernier, and Marceau (2009) also found a relationship between African-American achievement gap and SES and peer family structure. The common link in most of the research regarding family structure is that family influence appears to have the greatest impact during early childhood (Geoffroy et al., 2010). Suizzo, Robinson, and Pahlke (2008), in examining the influence of family on early learning, found that middle-class mothers struggled with approaching topics such as education and overcoming racial barriers with their children.

Sun and Li (2011) argued that it is not only family structure that has an impact on achievement, but also family stability. Sun and Li (2011) stated “decades of changing rates in marriage, divorce, remarriage, and cohabitation in the United States have led to two general trends in the living arrangements of American children” (p. 542). The researchers used a sample of more than 22,000 American kindergarteners to compare academic performance growth curves from kindergarten through fifth grade among three types of non-disrupted and three types of disrupted families. The results of the study indicated that students from non-disrupted, two-parent families made continuous academic progress while students from disruptive, step-parent households did not progress academically.

According to Clayton (2011), the actual school environment, in addition to racism and family structure, should also be considered major factors that impact academic achievement. However, Chaturvedi (2009) examined the interaction of a third variable, teacher quality, while controlling for SES as a confounding variable. In summary, Clayton (2011) found that students in high-poverty, high-minority, and low-performing schools have less access to well-qualified teachers. Harris and Sass’ (2011) research supported Clayton’s conclusions and added teacher experience and training produces larger achievement with experience. In addition, the results appeared to mirror the findings of other similar studies that indicated there were achievement

differences between African-American, White, and Hispanic students (Frankenberg, 2009; King & Bouchard, 2011).

Chaturvedi (2009) conducted a study to examine various school environmental factors and the relationship to achievement and motivation. Chaturvedi (2009), further concluded that the occupation of the father and type of school (i.e., private vs. public) were examined together with gender and grade. The results showed that all of the variables had a significant effect on academic achievement, demonstrating once again that school environment appears to be a factor in closing the achievement gap.

Success for All

Success for All (SFA) is one of the most utilized whole-school reading programs currently in use and has been shown to be successful based upon empirical research data (Institute of Education Sciences, 2007). According to Smith (2005), the SFA model is now used in approximately 1,300 schools in 47 states. Although the program was initially designed to address the reading proficiency of students in Grades K-3, it has since been expanded through the upper elementary and middle grades (Slavin et al., 2005).

Success for All (SFA) focuses on four major components: (a) 90-minute reading groups, (b) grouping by ability, (c) cooperative learning, and (d) continuous testing (Smith (2007). The overall effectiveness of the program has been well documented, utilizing the evaluation criteria of systematic and empirical, rigorous data analysis, reliable and valid data collection, strong research design, detailed results that allow for replication, and expert scrutiny (U.S. Department of Education, Comprehensive School Reform Program Office, 2002). In one program evaluation effort Slavin et al. (2005), utilized a longitudinal, quasi-experimental cluster randomized sample design over a two-year period to evaluate reading achievement. Schools were randomly assigned

to implement Success For All or control groups. Students from 38 schools were randomly selected and placed in pretest and posttest groups. The results indicated students in the SFA schools gained more than a half-year improvement in reading attack skills (Slavin et al., 2005).

In another evaluation study, Daniels, Madden, and Slavin (2004) examined the impact of SFA on middle school students. The program was evaluated by an external third party agency, the National Opinion Research Center. The research team utilized randomized pretest and posttest control groups to compare the impact of SFA on high-stakes reading achievement. Once again, the SFA students out gained the students in the non-SFA groups (Slavin et al., 2005).

The What Works Clearinghouse, Institute of Education Services, U.S. Department of Education (2009) and the Institute of Education Services (2007) “reviewed one hundred and ten studies on the impact of SFA in four areas: alphabetic, fluency, comprehensive, and general reading achievement” (p. 4). In general, the researchers rated SFA programs highly to moderately successful in most reading domains. None of the research studies found negative results. In effect, SFA equaled or exceeded all comparison groups’ performances.

NCLB (2001) identified specific subgroups, African Americans, and students from disadvantaged families as a central focus of the legislation. It is, therefore, absolutely necessary that any critical review of the effectiveness of any comprehensive school reform initiative be evaluated relative to their success with these two subgroups in particular. One such evaluation study was conducted by researchers at Johns Hopkins University and the College of Education at the University of Memphis (Borman et al., 2000). Although the purpose of the study was to understand the impact of selected whole-school reform models on high poverty Title I school students, the sample population was also comprised of a significant number of minority students, specifically African American and Hispanics. The whole-school reforms consisted of SFA,

Roots and Wings, Direct Instruction, and Core Knowledge. Nine schools participated in the mixed method research design. Four of the nine schools phased in Roots and Wings and Success for All over a 3-year period. In each school, the student population consisted of 99%-100% African-American students where 98% were eligible for free or reduced-price school lunch. The implementation of the reform models resulted in a number of mixed results. None of the whole-school reform models were a total success. However, there were qualitative results that included improved teacher perception of the reform effort (Slavin et al., 2005).

Based upon the evidence as reflected in the aforementioned studies (Borman et al., 2000; Daniels et al., 2004; Jackson, 2006; Slavin et al., 2005), it is apparent that Success for All is an effective whole-school reform model for improving reading proficiency of minority and disadvantaged students. However, school systems must consider the initial and on-going costs associated with adopting reform models like Success for All. According to the Success for All website (www.csoc.jhu.edu), the cost for first-year implementation of SFA is between \$70,000-\$85,000. Second-year program costs are between \$26,000 and \$30,000. Year three of the program costs between \$23,000 and \$25,000, resulting in a total cost of between \$119,000-\$140,000 (Borman & Hewes, 2002). Yet, Borman and Hawes (2002) concluded the following:

Success for All students completed eighth grade at a younger age, with better achievement outcomes, fewer special education placements, fewer retentions, and at the same educational expense. Further cost-effectiveness comparisons to the three prominent interventions suggest that Success for All is deserving of similar recognition as a sound educational investment that provides strong and lasting educational benefits. (pp. 243-266)

Direct Instruction

Direct Instruction, developed by Siegfried Engelmann and Wesley Becker of the University of Oregon, is among the oldest whole-school reform programs in the country. It was initially created to address the learning deficiencies of disadvantaged urban students (Adams & Engelmann, 2007; Engelmann, 1996; Maccoby & Zellner, 1970). Direct Instruction was also one of the initial sponsor programs for Project Follow Through, created under President Lyndon B. Johnson as an extension of the Head Start program and a part of the Great Society Programs created to offset the impact of poverty (Egbert, 1981). Direct Instruction consists of several specific components. Specially, instruction is highly scripted with a major emphasis on intense student-teacher interaction in which students and teachers receive immediate feedback. As reported by Stockard (2011), “The instructional approach also places great emphasis on the acceleration of the learning process” (p. 31). According to Huitt (2008), models of instruction are often based upon various theories of learning, such as operant conditioning and information processing. This purported connection between operant conditioning and Direct Instruction appears to be a contradiction in several respects (Englemann, 1968). Operant conditioning is based upon the belief that learning is shaped or contingent on the consequences of a particular behavior as opposed to the antecedents of behavior (Santrock, 2009; Woolfolk, 2010). If this is so, what are the specific consequences associated with direct instruction that would increase the probability that a student will repeat the behavior(s) that would lead to increased learning?

It would appear to the researcher that classical conditioning, where behavior is paired with a pleasant stimulus, may be more effective in ensuring that students repeat the positive antecedent behaviors that lead to increased learning. Further, according to several proponents of the Direct Instruction (Douglas, 2004; Huitt, 2008; Rosenshine, 1996; Santrock, 2009), the

Direct Instruction model is an interactive process between teacher and student. This being the case, it is logical to assume that the behavior of the teacher would also be a powerful antecedent impacting student learning. In view of the fact that operant conditioning, similar to direct instruction, is one such instructional model, and according to Rosenshine (1976), produces a greater rate on student learning as evidenced by improved student proficiency scores. Although direct instruction is often implemented as a whole-school reform model, the central focus is teacher behavior and student-teacher interaction (Rosenshine, 1996).

Although the literature provides several examples of the success of Direct Instruction (Huitt, 2008; Rosenshine, 1996; Santrock, 2009; Stockard, 2011), one of the initial evaluations of the reading program (Ryder, Sekulski, & Silberg, 2003) was highly critical by concluding that:

Across all the schools in the study, students in first, second, and third grade receiving Direct Instruction scored significantly lower on their overall reading achievement than students receiving more traditional forms of reading instruction and these results were consistent across 3 consecutive school years; and overall, on measures of reading achievement, students receiving more traditional forms of reading instruction in urban and suburban school districts display significantly greater gains than students receiving Direct Instruction. (p. 3)

Ryder et al. (2003) critical review of the Direct Instruction reading program resulted in a critical review of the evaluator's review (Adams et al., 2004). In fact, it is safe to say that the response to the criticism went beyond and above the typical response to a critical review. In addition to pointing out that Ryder's critique of DI was an unfair assessment of the process, the reviewers suggested that the entire critique of DI by Ryder should be disregarded because it was not published in a "peer-reviewed" journal (Tarver, as cited in Adams et al., 2004, p. 115).

The report contained three major flaws. The faulty conceptualization of Direct Instruction (which destroyed the integrity of the Direct Instruction that was being evaluated): The selection of so-called Direct Instruction classrooms in which the reading lessons were to be more like whole language or literature-based instruction than like real Direct Instruction lessons.... and grossly inadequate training of teachers in the purposes and use of Direct Instruction. After several meetings, I became convinced that Mr. Ryder's real intent was to provide so-called "evidence" that could be used to ridicule Direct Instruction. Otherwise, if he was a knowledgeable researcher, why would he propose a study so obviously flawed in both its design and implementation? (p. 115)

Historically speaking, this was not the first time that Direct Instruction was the target of a somewhat suspect evaluation of the effectiveness of the program. In 1977, a grant was issued by the Ford Foundation to evaluate Follow Through utilizing a third-party evaluation team (Egbert, 1981). The evaluators consisted of Gene Glass from the University of Colorado, Leslie McLean from the Ontario Institute for Education Studies, and Decker Walker of Stanford University. According to Glass (1981), "The truth about Follow Through is complex. No simple answer to the problem of educating disadvantaged students has been found...what worked well in one town worked poorly in another" (p. 4). Although Glass (1981) was evaluating Follow Through in general, in truth he and his team were evaluating all of the sponsored Follow Through programs, most specifically, Direct Instruction. In short, the evaluation, based upon the data presented, did not agree that highly structured programs were successful in improving achievement. Adams and Engelmann (1996), however, conducted a meta-analysis that indicated "mean effect size average per study was more than .75, which confirms that the overall effect is substantial" (p. 13).

a) Although literature abounds with individual and school testimonials to the effectiveness of Direct Instruction (Bock, Stebbins, & Proper, 1977; Carnine, Silbert, & Kameenui, 1997), one of the most compelling evaluation efforts comes from Engelmann (2007). Utilizing the data from the associates' evaluation, the authors provide evidence that Direct Instruction outperformed all of the Follow Through sponsor programs relative to its effectiveness with basic skills and disadvantaged populations (Becker & Engelmann, 1996). According to Becker and Engelmann (1996), the National Evaluation of Project Follow Through utilized a planned variation design to compare seven model programs relative to their overall effectiveness in teaching disadvantaged students. The evaluation involved 139 communities. The models compared were (a) The open Classroom Model (b) Cognitively-Oriented Curriculum Model (c) The Responsive Education Model (d) Bank Street Early Childhood Education Model (e) Tucson Early Education Model and (f) the Language Development (bilingual) Model.

The results of the comparative evaluation data showed that Direct Instruction outperformed all of the six models on grade equivalent and percentile performance scales by a quarter-standard deviation, a significant difference (Bereiter & Kurland, 1996).

Core Knowledge

The Core Knowledge educational reform movement is a national reading effort that has been shown to have positive results improving the reading ability of elementary level students and is most effective with urban students (Bradley, 2005; Dorn, n.d.; Stern, 2009). According to Hirsch (2008) “ Core Knowledge provides a clear outline of content to be learned grade by grade so that knowledge, language, and skills build cumulatively from year to year” (para. 2). Andrews (n.d.) provided further clarification of Hirsch's theory by adding that “sequenced-connecting

back with Hirsch's core belief that knowledge builds upon knowledge. Sequenced learning entails students' ability to acquire new knowledge by complementing what they already know” (para. 8).

Critics of the Core Knowledge curriculum, such as (Bradley, 2005; Dorn, n.d.; Stern, 2009, have suggested that individual differences are not addressed in the sequenced approach. From the perspective of the author, those who complain that Core Knowledge may not hold the interest of urban or minority students whose learning styles seem predicated on more current events may have a valid point. Other detractors (Dorn, n.d.) say that “The curriculum stresses Western culture and thereby shortchanges minority and low-income children” (para. 8).

Variables Impacting Reform Success/Failure

Burney et al. (2008) addressed the potential impact of extraneous variables on the possibility that whole-school reforms will either succeed or fail to succeed or fail to improve academic performance and thereby narrow or close the achievement gap. Foremost among these extraneous variables are SES, family structure, psychological factors and social and cultural factors SES, or more specifically poverty, has been the most researched and validated variable correlated with academic achievement (Yeung et al., 2008). It must be stated, however, that poverty or being disadvantaged is not a simple variable. In short, many of the other variables identified as having a possible interactive deleterious impact on achievement are also generally highly correlated with low SES or being disadvantaged (Burney & Beilke, 2008; Levin, 2007; Yeung & Conley, 2008). As a result, researchers and educators must be careful not to apply cause-effect relationships without extracting the potential interaction effect of other variables (Gall et al., 2003; Harris, 2011). As stated by Yeung and Conley (2008), “Other family

characteristics include family structure, number of children in the family, region of residence, and whether the family resided in a metropolitan area” (p. 310).

As another example of how variables outside of the context of the classroom may impact achievement, Oseguera, Conchas, and Mosqueda (2011) discussed how student achievement may be influenced by what they refer to as *social capital*. Social capital, as described by Oseguera et al. is described as “the less tangible resources gained through social relationships that positively influence educational outcomes” (p.1137). Oseguera et al. (2011) provided a descriptive explanation by stating that:

Unequal resources generate disparity in school engagement and achievement. These resources may be evident, such as financial support, school infrastructure, and technology, or less tangible, such as norms, encouragement, and information gained from relationships and social networks. (p. 1136)

One may conclude that social capital is very similar to the psychological construct of self-esteem that has been found to be a significant factor and correlate to academic achievement (Alloway, Gathercole, Kirkwood, & Elliott, 2009; Patten, 2011).

Summary

The review of the literature pertaining to NCLB (2001), while extensive, is yet mixed relative to drawing any conclusive impressions of the overall effectiveness of the NCLB. Historically, the core components of NCLB have a long and comprehensive story to convey. Having been created from a long list of previous national efforts to reform and improve education in general and kindergarten through Grade 12 achievement in particular, the legislation is comprehensive, yet filled with contradictions in both scope and process (Berends et al., 2002; Fullan, 2010). These same contradictions are evident when attempting to review the findings

relative to the overall effectiveness of the program. In short, the findings are mixed but appear to indicate that while there are success stories, the general feelings reported by educators in the field are that it either needs to be discarded or revamped.

The literature provides a fairly robust description of the effectiveness of various whole-school reform models. SFA is the most adopted and validated among those reform programs that are considered to be scientifically-based programs as described by the U.S. Department of Education. Direct Instruction has also been validated by numerous studies as being a scientifically-based effective reform program. Although Core Knowledge is not as well known or documented as Success for All or Direct Instruction, the literature contains a host of credible research studies, validating its place among scientifically-based whole-school reform programs.

The literatures examining the variables that consistently impact academic achievement appear to be in agreement on the fact that there is an achievement gap between the academic performance between African-American and White students. The achievement disparity is especially acute in reading and mathematics but shows in the other core subjects as well. The research is also consistent in pointing out that the achievement gap is not a recent or new phenomenon and has been evident and tracked since the early 1960s. As pointed out in the majority of the research, while there have been some recent improvements in academic achievement of both African Americans and Whites, the gap has remained constant.

A large body of the literature has addressed the wide range of causes associated with the achievement gap and generally agrees that many of the root causes, namely, poverty and culture, are elusive and difficult to correct. The earlier research reported in the literature, more so than current reports, alludes to biological more than social reasons for the achievement gap. Even the more conservative sources, such as the Heritage Foundation, now report that poverty and single-

parent households account for the achievement gap between African-American and White students. The literature is somewhat inconsistent regarding its reporting of the success or failure of the NCLB (2001) to achieve the primary purposes and goals. There are reports that NCLB generated significant gains in reading and mathematics. Upon closer examination, however, it becomes clear that the gains were generally concentrated in the White and Hispanic populations.

In summary, the achievement gap between African-American and White students is significant and appears to be widening in some instances. Achievement gaps are also present between and among other ethnicities; however, those gaps do not equal the differences between African Americans and Whites. Further, poverty is highly correlated with gaps in academic achievement within all ethnicities but is most prevalent within the African-American communities. Family structure is a significant factor and according to much of the literature plays a role in contributing to both levels of poverty and the achievement gap. In addition, race and racism are important factors when examining the origin of the achievement gap. Racism often produces a lasting psychological impact on African-American students. The literature is somewhat remiss in examining the impact of poverty on White academic achievement. The overall consensus of the literature is that while NCLB (2001) leaves much to be desired relative to the academic impact on kindergarten through Grade 12 education, there is general agreement that the intent of the Act was and remains appropriate and needed.

CHAPTER THREE: METHODOLOGY

NCLB (2001) mandates that school districts implement research-based reform programs to address deficiencies in reading among all students. Special emphasis is placed on this mandate relative to narrowing the achievement gap between minorities and other subgroups, such as disadvantaged students. More often than not, school districts employ expensive reform initiatives but seldom perform either formative or summative evaluations of these programs to justify continued utilization of same (Creswell, 2008; Wiersma & Jurs, 2005).

The purpose of the proposed study was to examine the impact of selected, whole-school, research-based instructional programs on African-American and disadvantaged students' CRCT scores in reading in a large urban school district. For purposes of this study, a disadvantaged student is one that qualifies for the respective school district's free or reduced-price school lunch program by meeting one or more of the Federal Government's Poverty Guidelines. These guidelines include but are not limited to income, number of persons in a household, or a member of a household that receives any form of public assistance. The income eligibility criteria is adjusted yearly (U.S. Census Bureau, 2011). Currently, a student who is a member of a household that meets the income criteria that is less than or equal to 185% of Federal Poverty Guidelines is classified as being "disadvantaged" and is therefore eligible for free or reduced-price school lunch. Local school districts are required to follow these guidelines (U.S. Department of Agriculture, 2011).

The methodology chapter includes a description of the setting for the study, the specific rationale used in the selection of the research setting, as well as a detailed description of the instrument and procedures used to identify and collect research data. The chapter also identifies

the research methodology and design of the study. The research questions and associated hypotheses provided an overview of the type of data analysis that is employed in the study.

Design

The researcher utilized a causal-comparative research design to examine the impact of three research-based educational reform programs on the reading achievement of African-American and disadvantaged students in Grades 3, 5, and 8. Causal-comparative research designs are appropriate for matching existing conditions with specific causes (Gay et al., 2009; Gall et al., 2003). In effect, causal-comparative is an “ex-post facto” approach for identifying reasons for the occurrence of a particular phenomenon, such as the impact of specific instructional approaches on academic proficiency (Lodico, Spaulding, & Voegtle, 2010).

Similar to other non-experimental research designs, causal-comparative and correlational designs seek to discover existing relationships. A significant difference is that causal-comparative designs attempt to attribute a cause-effect connection. As stated by Gall (2003), although causal-comparative research designs differ from correlational designs, the major difference is in how the variables are measured or analyzed (Gall, 2004, p. 296).

Threats to internal and external validity. According to Schenker, Phillip, and Rumrill (2004), experimental designs have a much stronger claim to internal validity than causal-comparative designs. Causal-comparative designs may or may not have a stronger claim to external validity, because external validity is established based on the degree to which the sample is representative of the larger population (Warner, 2008).

The researcher recognizes the fact that the lack of randomization is a potential threat to the overall validity of causal-comparative research and therefore utilized matching as a control variable to reduce the potential impact of extraneous variables that could result in committing a

Type I or II error (Gall et al., 2003; Gay et al., 2009). The researcher matched control variables based on socioeconomic status and ethnicity. A matching process was initiated to tease out factors that were not one of the grouping variables. As a result, students with disabilities, students who were not economically disadvantaged were not included in the analysis of the data.

Research Questions and Null Hypotheses

Research Question 1: Is there a significant difference in the school level third grade Reading performance between African American or disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African- American or disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho1: There is no significant difference in the school level third grade Reading performance between African American students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African American students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

Ho2. There is no significant difference in the school level third grade Reading performance between disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

Research Question 2. Is there a significant difference in the school level fifth grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core

Knowledge is implemented—and African-American students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho3. There is no significant difference in the school level fifth grade reading performance between African-American students in the intervention group and African-American students in the comparison group.

Ho4. There is no significant difference in the school level fifth grade Reading performance between disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

Research Question 3. Is there a significant difference in the school level eighth grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho5: There is no significant difference in the school level eight grade Reading performance between African American students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African American students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented

Ho6: There is no significant difference in the school level eight grade Reading performance between disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and

disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

Research Question 4. Is there a significant difference in the student level third grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American or disadvantaged students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho7. There is no significant difference in the student level third grade reading performance between African-American students in the intervention group and African-American students in the comparison group.

Ho8: There is no significant difference in the student level third grade reading performance between disadvantaged students in the intervention group and disadvantaged students in the comparison group.

Research Question 5. Is there a significant difference in the student level fifth grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and disadvantaged students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho9. There is no significant difference in the student level fifth grade reading performance between African-American students in the intervention group and African-American students in the comparison group.

Ho10. There is no significant difference in the student level fifth grade reading performance between disadvantaged students in the intervention group and African-American students in the comparison group.

Research Question 6. Is there a significant difference in the student level eight grade reading performance between African-American students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented

Ho11. There is no significant difference in the student level eighth grade reading performance between African-American students in the intervention group and African-American students in the comparison group.

Ho12. There is no significant difference in the student level eighth grade reading performance between disadvantaged students in the intervention group and disadvantaged students in the comparison group.

Research Question 7. Is there a significant difference between the relative impact of Success for All, Direct Instruction, or Core knowledge on the reading proficiency of students following a pairwise comparison between each of the reform models?

Ho13. There will be no significant difference between the relative impact of Success for All, Direct Instruction, or Core Knowledge on the reading proficiency of students following a pairwise comparison of each of the reform models.

Participants

The sample population for this study consists of African-American and disadvantaged students in grades 3, 5, and 8 from selected schools in a large, urban school system. For

purposes of this study, the 55 target schools were treated as the research subjects. The respective school reading performance constitutes the unit of analysis. The CRCT reading sub-test results includes the past year of 2010-2011. The participating schools previously self-selected one of the whole-school reform programs of Success for All (SFA), Direct Instruction (DI), or Core Knowledge. Each of the target schools presented a plan to the superintendent containing the respective whole-school reform program together with a rationale based upon their annual assessment results. Each school's selection was approved at the school level, the district level, and by the Superintendent.

The research population is made up of preexisting, intact groups; therefore, the targeted research participants will consist of aggregated CRCT reading test scores for students in grades 3, 5, and 8 in schools that implemented either Success for All, Direct Instruction or Core Knowledge whole-school reform programs. The targeted schools must have had one of the whole-school reform programs for a minimum of 4 years. In effect, schools selected as comparison schools will not have implemented any whole-school reform program. Control schools have a variety of instructional programs but have not implemented any whole-school reading reform programs.

The selection of schools representing the control group was matched on several criteria. The matching criteria are: similarity of student populations relative to achievement, ethnicity, and socioeconomic status (SES). A host of research studies have indicated that ethnicity and SES have historically been highly correlated with academic proficiency (Clewell et al., 2007; Garner, 2012). Research further indicates that SES in particular is highly associated with several factors that negatively impact what and how students learn (Strayhorn, 2010; Yeung & Conley, 2008). The other rationale for matching on the criteria of SES and ethnicity is that failure to match

schools' academic achievement, for example, increases the risk of committing Type I and II errors. In effect, matching on the aforementioned variables will assist in canceling out some extraneous factors. The sampling approach will be a cluster sampling technique because each sampling unit already exists and is intact (Wiersma & Jurs, 2005). Cluster sampling is appropriate in situations where the population members are naturally grouped in units (Creswell, 2007).

The researcher conducted a power analysis for the purpose of determining the appropriate sample size to minimize the chances of committing a Type I error. The results of the power analysis indicated that with a predetermined error rate of 5%, a confidence level of .95 and a unit of analysis population consisting of 61 schools, a sample size of 55 data points (schools) would be required. The actual research sample consisted of 61 schools. Table 1 provides specific demographic data for the district and student analysis.

Table 1

District and School Level Demographics

<u>District Level</u>					
District Schools	Required 95% CI	Sample size	Number tested	% Disadvantaged	% African American
78	55	61	11,459	75%	92%

Grade	<u>African-American students</u>		<u>Disadvantaged students</u>	
	Reforms	No Reforms	Reforms	No Reforms
3	1425	1260	970	1302
5	966	1171	1081	1885
8	1142	1157	1122	1252
Totals	3533	3594	3110	4440

Setting

The setting for the research was 55 elementary schools and 16 middle schools in a large urban school system. The student population of the selected school system consisted of approximately 51,000 students with 90% being African American, 6% White, and 2% Hispanic. Approximately 70% of the student population is eligible for free or reduced-price school lunch. The research setting was chosen primarily because it is the only system in the metropolitan area where adopting an educational reform to meet the NCLB requirement is mandatory based upon individual achievement goals.

Instrumentation

The Georgia Criterion Reference Competency Test (CRCT) was the instrument used to measure the reading proficiency scores for African-American and disadvantaged students in Grades 3, 5, and 8. The CRCT is designed to provide information and quantitative data that provides a profile of how well students in Grades 1 through 8 acquire and learn the skills set forth in Georgia's curriculum content as outlined in Georgia's performance standards (Georgia Department of Education, 2010). According to the GA DOE Assessment Division (2012):

The number of CRCT test items varies by year, subject, and grade level. The information collected from the test is used to assess academic achievement at the student, class, school, system, and state levels; additionally, the test is used to identify areas of individual student strengths and weaknesses. Raw scores are converted to scale scores for reporting of performance levels. Three levels of performance are used: Does Not Meet Expectation (below 800), Meets Expectations (800-849), and Exceeds Expectations (850 or above). (p. 2)

The CRCT is a criterion-referenced test specific to standards in Georgia; therefore, it is not a norm-referenced test. The goal of norm-referenced tests is to rank and sort students rather than measure the mastery of a subject (Georgia Department of Education, 2011). A criterion-referenced test is used to measure a student's mastery of a predetermined criteria or standard (Bond, 1996; Huitt, 1996; Popham, 1975). According to Bond (1996a):

In 1994, thirty-one states administered norm-referenced tests (NRTs), while 33 states administered criterion-referenced tests. Among these states, 22 administered both. Only two states rely on NRTs exclusively, while one state relies exclusively on a CRT. (p. 2)

More important to the purpose of this study is that criterion-referenced tests are designed to find out how much students know before instruction begins and after it has finished (Georgia Department of Education, 2011). For purposes of this study, the CRCT results, reported as percentages, were compared to the impact of selected whole-school reading reform programs on the percentage of students meeting or exceeding pre-established reading performance standards assessed in science and social studies. The CRCT only assesses the content standards outlined in the Georgia Professional Standards (Georgia Professional Standards Commission, 2009; Porter, 2007).

For the purposes of this study, Levels 2 and 3 (Meets and Exceeds) will be aggregated to represent the percentage of students who meet the achievement standards established by the Georgia Performance Standards (GPS). The GA DOE oversees the development of the CRCT and adheres to the Standards for Educational and Psychological Testing (1999) as established by the American Educational Research Association (AERA), the APA, and the National Council on Measurement in Education (NCME) (Georgia Department of Education, 2011).

Validity

Validity is one of the two cornerstones of technical quality in testing and measurement, which begins with the purpose of the assessment and continues through item writing and review (Georgia Department of Education (2010)). All CRCT items are written by qualified, professional content specialists specifically for the Georgia CRCT. The professional content specialists are selected from local school districts and institutions of higher learning. Classroom teachers, curriculum specialists, administrators, State Department of Education content specialists and education professors from colleges and universities throughout the state are included in the process. After the items are written, curriculum specialists and committees of Georgia educators

review the items. Items are evaluated for overall quality and clarity, content coverage and appropriateness, alignment to the curriculum, and grade-appropriate stimuli with an emphasis on higher-order thinking skills (Georgia Department of Education, 2007). Great care is taken throughout the item development process to monitor items for potential bias and to ensure representation of all Georgia's students. The GA DOE ensures that the CRCT meets the highest standards of technical quality and defensibility. The Testing Division meets with an independent panel of experts—Georgia's Technical Advisory Committee (TAC)—on a quarterly basis. TAC members are experts in the field of educational measurement who review all aspects of the test development and implementation process on a continual basis (GA DOE, 2007).

The identification for the purpose of any test, as described by the GA DOE (2007), is the first step in the validation process. The Georgia state legislature identifies the test purpose as creating a measure of how well students have mastered the State's curriculum (GA DOE, 2011b). The next step in the CRCT validation process is to form representative committees of educators to identify skills, knowledge, and concepts to be assessed.

One of the most important functions of the respective validation committees is to develop content domain specifications. These specifications are converted into content descriptions that are shared with all interested stakeholders. These documents become a part of the paper trail as evidence of the CRCT's validity as a measure of the State curriculum. One of the final steps in the CRCT validation process is the item writing phase. CRCT test items are written by qualified committees of Georgia educators and assessment specialists and aligned with the Georgia curriculum. Finally, the items are field tested for content and construct consistency. Standards are established following the first administration of the field test. Although this standardization phase may appear to be similar to what is typically called *norming*, it is not. The standards are

state-specific. The standard setting in this instance is where the committees decide the test items and the number of correct answers required to meet or exceed expectations. The last phase of the CRCT development is the conversion of raw scores into scaled scores and performance levels. The assessment department contracts with external, independent agencies to conduct alignment studies to ensure content validity. Finally, the department has collected evidence through separate independent alignment studies to ensure that the test measures the State's curriculum. The CRCT department also conducts analyses as evidence of external validity by comparing how the constructs the CRCT measures compare with other well-recognized assessments, for example, the Iowa Tests of Basic Skills (ITBS).

Reliability

The other cornerstone of technical quality in testing and measurements is reliability (GA DOE, 2010). In simple terms, reliability asks will the same measurement give the same or comparable result for the same student every time it is administered. Reliability also refers to the ability to generalize scores over time (test-retest), between other scores (interrater), and with other items within the test (internal consistency) (Gall, Gall, & Borg, 2004). Reliability is a means for determining standard error of measurement and constructing confidence intervals to identify where a person's true score is likely located within a specified range of scores (Salvia & Ysseldyke, 2007). The 2004 CRCT total test reliability for 2011 ranged from 0.89 to 0.90 for third grade Reading, 0.85 to 0.91 for English/Language Arts, and 0.87 to 0.91 for Mathematics. The GA DOE is confident that the CRCT are both reliable and valid (Georgia Department of Education, 2007). According to the GA DOE Standards (2011):

Reliability is the degree to which test scores for a group of test takers are consistent over repeated applications of a measurement procedure and hence are inferred to be

dependable, and repeatable for an individual test taker; the degree to which scores are free of errors of measurement for a given group. (p. 4)

The GA DOE reports several indices of reliability for the CRCT (Table 2). Cronbach's alpha reliability coefficient is perhaps the most well-known and the most robust (Cronbach, 1951). According to GA DOE (2011b), a reliability coefficient expresses the consistency of test scores as the ratio of true score variance to observed total score variance (i.e., true score variance plus error variance). The Georgia State Department of Education utilizes Cronbach's alpha as a measure of the internal consistency of the CRCT. Cronbach's alpha is computed using Crocker and Algina's formula (1986).

The Georgia assessment unit uses the standard error of measurement (SEM) as the second index of reliability for the CRCT Georgia Department of Education (2007). This index addresses the random variability of the raw test scores. The following table shows the reliability indices in terms of Cronbach's alpha along with the raw score SEM for all grades and subjects of the 2011 CRCT.

Table 2

Reliability Coefficients (Cronbach's Alpha) and Raw Score SEM for Subject/Tests by Grade

Grade	<u>Reading</u>	
	SEM	Alpha
3	.89	2.45
4	.90	2.36
5	.88	2.53
6	.87	2.40
7	.87	2.51
8	.85	2.33

Note. From Georgia Department of Education (2011).

In general, the reliability coefficients for the CRCT for Grades 3, 5, and 8 are strong and support the contention that the CRCT provides consistent results across grade levels.

Procedures

The initial procedures addressed were obtaining the required Internal Review Board (IRB) packet and acquiring approval from the appropriate official. Following the IRB approval, the researcher inquired about any school district approval that might be needed; however, in view of the fact that the study only utilized public domain data and no human subjects were to be involved, system approval was not necessary. Using the public domain State Department of Education website, the researcher obtained a list of all schools in the urban school district to be included in the study. The list was verified and grouped into elementary and middle schools groups respectively. Information was obtained from the selected school systems; research department to verify the reform model implemented in each school relative to the respective reform model and the period of implementation. Each of the respective school reform models were matched with each of the schools, including the number of years that the reform model had been in place.

The Georgia State Department of Education, Department of Assessment and Accountability provided data files for each of the respective schools included in the study to verify GRCT scale scores and percentage of students in each performance level.

Data Analysis

The dependent variable was Reading performance at the third, fifth, and eighth grade levels, respectively. Reading performance is categorical or dichotomous in that students will be classified into two groups: meeting/exceeding established performance levels in reading on the *Georgia Criterion Referenced Competency Tests (GCRCT)* or not meeting/exceeding established performance levels in reading on the *Georgia Criterion Referenced Competency Tests (GCRCT)*.

There is one independent variable group with two levels: intervention and comparison. Data will be entered into SPSS, a statistical package for the data analyses. These data have intervention and comparison group students in grades 3, 5, and 8, respectively. The groups were coded 1 for intervention group and 0 for comparison group. There is one dependent variable with two levels or categories: Meets/exceeds performance standards and **does** not meet/exceed performance standards. Those students whose scores meet or exceed performance expectations were coded, one (1). Those who do not meet or exceed performance expectations were coded zero (0). Therefore, an independent-sample chi square test was appropriate for testing the null hypotheses. As Huck (2000) states, when two independent samples are compared with respect to a dichotomous dependent variable, the chi-square test can be thought of as analogous to an independent samples *t*-test (p. 619). The null hypotheses for this form of a chi-square test contends that the two populations are identical in the percentage split between the two categories of the response variable. In this dissertation research, the proportions of students exposed to the intervention are hypothesized to meet/exceed reading performance expectations to a greater extent than those students in the comparison group who do not have the benefit of the intervention/treatment. A probability level $p < .05$ was used to establish statistical significance.

Each of the thirteen null hypotheses was tested by comparing the differences between the percent of schools scoring at the aggregated meet and exceeds levels on the reading subtest of the CRCT. The data will be analyzed for African-American and disadvantaged students in Grades 3, 5, and 8 for the 2010-2011 school year comparisons were made between schools meeting/exceeding reading performance in each of the whole-school reform schools and students where no reform was implemented.

An independent samples chi-square test is a nonparametric test was utilized to detect any significant differences between the reading proficiency of each type of education reform (Gall et al., 2003; Gay, Mills, & Mills, 2009; Thompson, 2002). Independent samples Chi-square tests are appropriate for determining whether or not the means of several dependent variables are equal (Gall et al., 2003). The chi-square, as opposed to multiple t-tests, also decreases the chances of committing a Type I error (Creswell, 2003; Gall et al., 2003; Gay et al., 2009). For purposes of the proposed research study, the approach permitted the researcher to measure the differences between the percent of students performing at the meet or exceed levels (both levels were aggregated to form one statistic) of the CRCT in reading for each of the three research-based reform models. The second phase of the analysis addressed the second part of the purpose of the study, namely to determine which of the three reform models resulted in the greatest reading proficiency gain by comparing each of the reform models with each other for each grade level.

Effect size is a process used to provide a deeper and more comprehensive meaning and understanding about the magnitude of the difference between two groups following a statistical analysis such as analysis of variance (ANOVA) (Gall et al., 2003; Gay, Mills, & Mills, 2009; Thompson, 2002) or a Chi-square test (Aguinis, 2006; Grissom & Kim, 2005). According to the American Psychological Association (APA) guidelines (2009), the results section of all research should include some measure of the effect size to assist the reader to appreciate the magnitude of the differences including some discussion about the confidence interval associated with each effect size (Huck, 2012).

In practical terms and for purposes of this research, *effect size* serves the purpose of quantifying any differences between the impact of each of the independent variables (Success for

All, Direct Instruction, and Core Knowledge) on the reading proficiency of African-American and disadvantaged students in Grades 3, 5, and 8. Effect size calculations are especially useful for evaluating the degree to which respective educational interventions, such as whole-school reading reforms, impact the percentage of students meeting the reading standards established by a particular school district (Warner, 2008). Effect size is also important due to the fact that practical significance of the findings must be considered. For purposes of this study, the size of the differences between the impact of the respective reading reforms are significant; however small, school districts will want to compare academic benefits against the practical considerations of costs and resources. For purposes of this study, Cramer's V —considered by several statisticians to be the most reported effect size for chi-square—will be utilized as the effect size statistic for this study (Grissom & Kim, 2012; Warner, 2008). Cramer's V is also appropriate for chi-square analysis involving more than two rows or columns. As recommended by Warner (2008), analysis of any chi-square results should include a table showing frequencies of columns and row percentages; obtained value of chi-square and the degrees of freedom; and a brief discussion about the nature of the relationship reported in probabilities. A Mann-Whitney U test was used to conduct the pair-wise comparison between the three reform models. The Mann-Whitney test is similar to other tests that compares two independent samples. The Wilcoxon and Levene's median tests are similar but are stronger than the Mann-Whitney (Huck, 2012).

CHAPTER FOUR: FINDINGS

Analysis Procedures

This quantitative, causal-comparative research study has two purposes. The initial purpose was to conduct an in-depth examination of the degree to which selected whole-school reading reform programs could potentially affect the reading proficiency of students in Grades 3, 5, and 8 in a large urban school district. More to the point, the primary purpose of this examination was to focus mainly on the degree to which whole-school reading reform programs impacted the reading achievement of the underserved groups of African-American and disadvantaged students. For purposes of clarity, the impact of the whole-school reading reforms was measured by the percentage of students whose reading proficiency met or exceeded the reading standards as evidenced by the Criterion Referenced Competency Test (CRCT). Success for All, Direct Instruction, and Core Knowledge were the whole-school reading reforms constituted the independent treatment variables. One of the general overriding issues of this study was to provide quantitative evidence addressing whether the identified reading reforms impacted reading proficiency across the grade levels of 3, 5, and 8 for either African-American and/or disadvantaged students.

The second purpose of the study was to determine which of the three whole-school reading reform programs produced the greatest percentage of students scoring in the meet or exceed reading levels on the CRCT for African-American and disadvantaged students in grades 3, 5, and 8. This question was addressed by conducting a pairwise comparison between the three whole-school reforms. In effect, which reading reform programs, without consideration of the control variables of ethnicity, socioeconomic status, or grade level produced the highest reading proficiency? The entire study and analysis is guided by the following research questions:

Research Question 1

Is there a significant difference in the school level third grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Research Question 2

Is there a significant difference in the school level fifth grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Research Question 3

Is there a significant difference in the school level eight grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American or disadvantaged students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Research Question 4

Is there a significant difference in the student level third grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is

implemented—and African-American or disadvantaged students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Research Question 5

Is there a significant difference in the student level fifth grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—African-American and disadvantaged students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?.

Research Question 6

Is there a significant difference in the student level eight grade reading performance between African-American students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Research Question 7

Is there a significant difference between the relative impact of Success for All, Direct Instruction, or Core knowledge on the reading proficiency of African-American or disadvantaged students following a pairwise comparison between each of the reform models?

The analysis of results consisted of three phases. The first phase of the analysis examined and compared schools that had implemented whole-school reading reform programs to schools that had not implemented whole-school reform programs relative to the level of reading proficiency of African-American and disadvantaged students as evidenced by the percentage meeting or exceeding State reading standards. This phase of the analysis focus was at the school

level. The second phase of the analysis examined and compared differences in the percentage of students attaining reading proficiency at the individual student level in grades 3, 5, and 8.

The former set of analyses may be of particular interest to school administrators who wish to know whether a set of reform practices will increase their standing relative to other school buildings in terms of the percentage of students attaining proficiency. However, the student-level analyses have higher levels of statistical power, because there the analysis is based on a large number of students, rather than a smaller set of schools. At both the school and the student level, separate analyses examined the percent of proficiency attained for African-American students, and for disadvantaged students, at the third, fifth, and eighth grades.

The third stage of the analysis is designed to provide quantitative information relative to the specific impact of each of the whole-school reading reforms by conducting pairwise comparisons between each of the three reform models in terms of improving proficiency rates for all students. The reform models compared include: Success for All, Direct Instruction, and Core Knowledge.

School-Level Analysis

In order to determine whether there were significant differences in the percent proficient between reform and non-reformed schools, a Mann-Whitney U test was employed. The Mann-Whitney U test was utilized, rather than an independent groups t-test, because the distribution of percent proficient across schools does not follow a normal distribution, and because the sample sizes of schools, particularly at grade 8, is relatively small (Huck, 2012). The Mann-Whitney U test assessed the null hypothesis that the median proficiency rate for schools is equal for the reformed and non-reformed schools. The null hypothesis is evaluated by comparing the median rank of schools in the reformed group and the non-reformed school in terms of the school

proficiency rate. To the extent that the median rank of schools in one group is different from the other group, then it is less likely that the sample was drawn from a population in which the Null Hypothesis was true. Where the Null Hypothesis has been rejected, the effect size for the reformed versus non-reformed comparison was computed, following procedures described by Grissom and King (2012). Specifically, the probability that a random case drawn from the population in one group will be higher than a random case drawn from the other. This probability is computed by dividing the Mann-Whitney U test by the product of the sample sizes.

School-Level Results Grade 3

Research Question 1: Is there a significant difference in the school level third grade Reading performance between African American or disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African- American or disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

***H₀₁*:** There is no significant difference in the school level third grade Reading performance between African American students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African American students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

As shown in Table 3, no statistically significant differences in third grade reading proficiency rates for African American ($U=203.50$, $p=.365$). As a result, the researcher failed to reject the H_{01} .

***H₀₂*.** There is no significant difference in the school level third grade Reading performance between disadvantaged students in the Intervention Group where the whole-school

reform of Success for All, Direct Instruction, and Core Knowledge is implemented disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

The analysis of Ho2 that examined the impact of whole-school reforms on the reading proficiency of disadvantaged (Table 3) also resulted in a failure to reject the hypothesis of no difference (Mann-Whitney U of 128, $p=.155$).

Table 3

Grade 3 Proficiency Rates for Reform and Non-reform Schools (School Level Analysis)

		African-American			
School type	n	Mean percent proficient	Median rank	Mann-Whitney U statistic	p
Non-reform	22	85.50	24.25	203.50	.365
Reform	17	82.32	20.75		
		Disadvantaged			
School type	n	Mean percent proficient	Median rank	Mann-Whitney U statistic	p
Non-reform	22	85.82	17.32	128.00	.155
Reform	17	89.56	22.50		

School-Level Results Grade 5

Research Question 2. Is there a significant difference in the school level fifth grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho3. There is no significant difference in the school level fifth grade reading performance between African-American students in the intervention group and African-American students in the comparison group.

As shown in Table 4, no statistically-significant differences in fifth grade reading proficiency rates for African-American ($U=185, p=.955$). The results of the Mann-whitney U caused the researcher to fail to reject $H_0 3$.

Ho4. There is no significant difference in the school level fifth grade Reading performance between disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

Again, Table 4 indicates that like African-American students, the reading proficiency of disadvantaged students was not significantly impacted by reform or non-reform school programs resulting in a failure to reject H_04 ($U=143.50, p=.336$).

Table 4

Grade 5 Proficiency Rates for Reform and Non-reform Schools (School Level Analysis)

		African-American			
School type	n	Mean percent proficient	Median rank	Mann-Whitney <i>U</i> statistic	<i>p</i>
Non-reform	22	86.91	19.91	185.00	.955
Reform	17	87.47	20.12		
		Disadvantaged			
School type	n	Mean percent proficient	Median rank	Mann-Whitney <i>U</i> statistic	<i>p</i>
Non-reform	22	86.00	18.02	143.50	.336
Reform	17	88.13	21.53		

School-Level Results Grade 8

Research Question 3. Is there a significant difference in the school level eighth grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American or disadvantaged students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho5: There is no significant difference in the school level eight grade Reading performance between African American students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and

African American students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

As shown in Table 5, no statistically-significant differences in eight grade proficiency rates for disadvantaged students ($U=29.50, p=.329$) The results of the Mann-Whitney U caused the researcher to fail to reject H_05 . In short, African-American students in grade five did not demonstrate any significant differences in reading proficiency when attending reform or non-reform schools.

H₀₆: There is no significant difference in the school level eight grade Reading performance between disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

The results of the Mann-Whitney U of 36.50 and a probability of .723 caused the researcher to fail to reject H_06 indicating that at the eighth grade level, the reading proficiency of disadvantaged students was not significantly different (Table 5).

Table 5

Grade 8 Proficiency Rates for Reform and Non-Reform Schools (School Level Analysis)

		African-American			
School type	<i>n</i>	Mean percent proficient	Median rank	Mann-Whitney <i>U</i> statistic	<i>p</i>
Non-reform	5	94.44	10.72	29.50	.329
Reform	11	91.11	8.28		
		Disadvantaged			
School type	<i>n</i>	Mean percent proficient	Median rank	Mann-Whitney <i>U</i> statistic	<i>p</i>
Non-reform	5	91.67	9.06	36.50	.723
Reform	11	94.33	9.94		

Student-Level Analysis Findings for Grades 3, 5, 8

In order to determine whether there were significant differences in proficiency rates between African-American and disadvantaged students in reformed and non-reformed schools, Pearson's chi-square test was employed. The chi-square test assesses the null hypothesis that the proportion of students attaining proficiency is equal for students in reformed and non-reformed schools. This test assumes that cases can be categorized into mutually exclusive categories (proficient and non-proficient). Although an entire school cannot be categorized as being proficient or not proficient, and individual student can be. When significant differences were found, effect sizes were estimated using the Phi coefficient. Phi can be interpreted in a manner similar to a coefficient of correlation (Huck, 2012; Warner, 2008).

Student-Level Findings Grade 3

Research Question 4. Is there a significant difference in the student level third grade reading performance between African-American or disadvantaged students in the intervention

group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American or disadvantaged students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho7. There is no significant difference in the student level third grade reading performance between African-American students in the intervention group and African-American students in the comparison group.

As shown in Table 6, proficiency rates were significantly higher (chi-square = 4.461, $p = .035$) among African-American students in non-reformed schools (85.78%) than among African-American students in reformed schools (82.81%). The results of the chi-square analysis resulted in rejection of the hypothesis of no difference indicating that African-American students performed significantly better in reading than did African-American students who attended schools implementing whole-school reading reforms. Although the differences were statistically significant, the Phi coefficient effect size (.04) indicates a weak relationship the independent and dependent variables.

Ho8: There is no significant difference in the student level third grade reading performance between disadvantaged students in the intervention group and disadvantaged students in the comparison group.

By contrast, Table 6, proficiency rates were also significantly higher (chi-squared = 4.833, $p = .028$) among disadvantaged students in reformed schools (89.48%) compared with disadvantaged students in non-reformed schools (86.43%). The Phi coefficients (effect sizes) for the reformed versus non-reformed models was .05 for disadvantaged students. The results of the chi square

resulted in a rejection of H_0 . Please note that a very small effect can attain statistical significance in a sample of this size.

Table 6

Grade 3 Proficiency Rates for Reform and Non-Reform Schools (Student Level Analysis)

School type	African-American			χ^2	p
	Pass	Fail	Pass %		
Non-reform	1086	180	85.78	4.461	.035
Reform	1180	245	82.81		

$\chi^2 (2, N=2,690) = 4.461, p < .028$

School type	Disadvantaged			χ^2	p
	Pass	Fail	Pass %		
Non-reform	1127	177	86.43	4.833	.028
Reform	868	102	89.48		

$\chi^2 (2, N = 2,274) = 4.833, p < .028$

Student-Level Findings Grade 5

Research Question 5. Is there a significant difference in the student level fifth grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and disadvantaged students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

H_0 9. There is no significant difference in the student level fifth grade reading performance between African-American students in the intervention group and African-American students in the comparison group.

As shown in Table 7, no significant differences were found in proficiency rates between fifth grade African-American students in reformed and non-reformed schools ($\chi^2=.034, p=.854$).

The chi-square statistic results caused the researcher to fail to reject H_09 : In effect, African-American students reading proficiency did not vary significantly between those attending schools with reading reforms and those not implementing reforms.

H₀₁₀. There is no significant difference in the student level fifth grade reading performance between disadvantaged students in the intervention group and African-American students in the comparison group.

Similarly, no significant differences were found in proficiency rates between disadvantaged students in reformed and non-reformed schools in Grade 5 ($\chi^2= 2.832, p=.092$). As a result of the chi square findings, the researcher failed to reject H_010 .

Table 7

Grade 5 Proficiency Rates for Reform and Non-Reform Schools (Student Level Analysis)

School Type	African-American			χ^2	p
	Pass	Fail	Pass %		
Non-reform	1019	152	87.02	.034	.854
Reform	838	128	86.75		

$$\chi^2(2, N = 1,856) = .034, p < .54$$

School type	Disadvantaged			χ^2	p
	Pass	Fail	Pass %		
Non-Reform	1621	264	85.99	2.832	.092
Reform	898	120	88.21		

$$\chi^2(2, N = 2,127) = 2.518, p < .092$$

Research Question 6. Is there a significant difference in the student level eighth grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American or disadvantaged students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Student-Level Findings Grade 8

Ho11. There is no significant difference in the student level eighth grade reading performance between African-American students in the intervention group and African-American students in the comparison group.

As shown in Table 8, at the eighth grade level, no significant differences ($\chi^2=3.674$, $p=.055$) were found in proficiency rates between African-American students in reformed and non-reformed schools. As a result, the researcher failed to reject the hypothesis of no difference for Ho11 indicating that the reading proficiency of eighth grade African-American students resulted in no significant differences based the type of reading program being implemented.

Ho12. There is no significant difference in the student level eighth grade reading performance between disadvantaged students in the intervention group and disadvantaged students in the comparison group.

As shown in Table 8 The chi-square results, $\chi^2= 8.347$, was significant, $p = .004$, resulting a the rejection of the hypothesis of no difference, indicating that the reading proficiency of disadvantaged students attending schools with whole-school reading reform programs outperformed disadvantaged students attending non-reform schools. The strength of the relationship between the independent variable and the outcome variable was small, .06.

Table 8

Grade 8 Proficiency Rates for Reform and Non-Reform Schools (Student-Level Analysis)

School type	African-American			χ^2	<i>p</i>
	Pass	Fail	Pass %		
Non-reform	1083	74	93.60	3.674	.055
Reform	1045	97	91.51		

$$\chi^2(2, N = 2,127) = 3.674, p < .055$$

School type	Disadvantaged			χ^2	<i>p</i>
	Pass	Fail	Pass %		
Non-reform	1150	102	91.85	8.347	.004
Reform	1064	58	94.83		

$$\chi^2(2, N = 2,213) = 8.347, p < .004$$

Comparison of Reform Models

The schools implementing whole-school reforms in the sample followed one of three models: Success for All (SFA), Direct Instruction (DI), and Core Knowledge (CK). The next stage of the analysis compared proficiency rates between each of the three models at the school level in an effort to rate the effectiveness of each respective model in improving the reading proficiency of the general population of African-American and disadvantaged students across Grades 3, 5, and 8.

Whole-School Comparison Findings

Research Question 7. Is there a significant difference between the relative impact of Success for All, Direct Instruction, or Core knowledge on the reading proficiency of students following a pairwise comparison between each of the reform models?

Ho13. There will be no significant difference between the relative impact of Success for All, Direct Instruction, or Core Knowledge on the reading proficiency of students following a pairwise comparison of each of the reform models.

Direct Instruction (DI) versus Success for All (SFA). As shown in Table 9, the percentage of students attaining proficiency was significantly higher ($U = 33.50, p < .001$) in schools implementing the Direct Instruction reform model ($M = 88.34$) compared with those that were implementing the Success for All reform model ($M = 81.81$). The probability that a randomly drawn score from the Direct Instruction population will be greater than one drawn from the Success for All population is .13, resulting in a rejection of Hypothesis 13, indicating that there was a significant difference between the impact of Direct Instruction and Success for All on reading proficiency across grade levels, ethnicity, and socioeconomic status.

Table 9

Proficiency Rates for (DI) Schools and (SFA) Schools

Reform model	Mean percent proficient	Median rank	Mann-Whitney U statistic	p
DI	88.34	22.41	33.50	.001
SFA	81.81	10.59		

$U = 33.50, p < .001$

Core Knowledge (CK) versus Success for All (SFA). As shown in Table 10, the percentage of students attaining proficiency was significantly higher ($U = 11.00, p < .001$) in schools implementing the Core Knowledge reform model ($M = 91.44$) compared with the Success for All reform model ($M = 81.81$). The probability that a randomly-drawn score from the Core Knowledge population will be greater than one drawn from the Success for All

population is .13. The results of the Mann-Whitney were a rejection of H_0 , indicating that there was a significant difference between the impact of Core Knowledge and Success for All on reading proficiency across grade levels, races/ethnicities, and socioeconomic statuses.

Table 10
Proficiency Rates for (CK) and (SFA) Schools

Reform model	Mean percent proficient	Median rank	Mann-Whitney <i>U</i> statistic	<i>p</i>
CK	91.44	24.89	11.00	.001
SFA	81.81	9.19		

$U = 11.00, p < .001$

Core Knowledge (CK) versus Direct Instruction (DI). As shown in Table 11, differences in proficiency rates between schools implementing the Core Knowledge and Direct Instruction reform models are not statistically significant. The probability that a randomly drawn score from the Direct Instruction population will be greater than one drawn from the Core Knowledge population is .04.

As a result of the Mann-Whitney *U* statistic, the researcher failed to reject the hypothesis of no difference, indicating that there was no significant difference between the impact of Core Knowledge and Direct Instruction on reading proficiency across grade levels, ethnicity, and socioeconomic status.

Table 11

Proficiency Rates for (CK) and (DI) Schools

Reform model	Mean percent proficient	Median rank	Mann-Whitney <i>U</i> statistic	<i>p</i>
CK	91.44	20.25	94.50	.087
DI	88.38	14.41		

$U = 94.50, p > .087$

Summary

The researcher utilized a causal-comparative research design to examine the impact of three research-based educational reform programs on the reading achievement of African-American and disadvantaged students in Grades 3, 5, and 8. An independent-sample chi-square test was used to evaluate differences between the percent of schools scoring at the aggregated meet and exceed levels on the reading subtest of the CRCT. The Mann-Whitney *U* test was applied for the purpose of generating the effect size of each comparison. The analyses consisted of school-level and student-level analyses to address the seven research questions and null hypotheses. The analysis also included a comparison between the effectiveness of each of the respective whole-school reading reforms.

The school level analyses did not provide any evidence of differences in student achievement between the reformed and non-reformed schools. At the individual student level, significantly higher proficiency rates were found for disadvantaged students in grade 3 and grade 8 in reformed compared with non-reformed schools. Among African-American students, significantly higher proficiency rates were found in grade 3 for students in non-reformed schools,

compared with reformed schools. The differences in results between the school-level and student-level analyses are due to the greater sensitivity afforded by analyzing data from hundreds of students. Proficiency rates differed significantly between reform models. Significantly higher proficiency rates were found in schools implementing the Direct Instruction or the Core Knowledge reform models compared with the Success for All model.

CHAPTER FIVE: SUMMARY AND DISCUSSION

The sample population for this study consisted of CRCT reading proficiency scores of students from 61 schools in Grades 3, 5, and 8 who took the CRCT reading test. The number of African-American students who took the test in schools where whole-school reform programs were implemented totaled 3,533, compared to 3,594 students taking the test from students attending schools where no whole-school reform programs were implemented. The number of disadvantaged students that took the CRCT reading test was 3,110 in schools with whole-school reforms and 4,440 who attended schools that did not implement reforms (Table 11).

Table 12

Sample Population of African-American and Disadvantaged Students in Grades 3, 5, and 8

Grade	African-American students		Disadvantaged students	
	Reforms	No Reforms	Reforms	No Reforms
3	1425	1260	970	1302
5	966	1171	1081	1885
8	1142	1157	1122	1252
Totals	3533	3594	3110	4440

This chapter addresses each of the seven research questions by quantitatively examining results of the null hypotheses utilizing the chi square and the Mann-Whitney *U* statistic. Effect size results are provided in instances where the null hypothesis of no difference is rejected in order to provide a quantifiable understanding of the relative strength of the differences between groups. Success for All (SFA), Direct Instruction (DI), and Core Knowledge (CK) on the reading

proficiency of African-American and disadvantaged students in Grades 3, 5, and 8 as evidenced by the percentage scoring in the *Meet* and *Exceed* levels of the Criterion Referenced Competency Test (CRCT).

Summary of Findings and Discussion

The primary purpose of the study was to examine the degree to which three selected scientifically-researched whole-school reforms impacted the reading proficiency of African-American and disadvantaged students. Reading proficiency in Grades 3, 5, and 8 was measured by the percentage of schools and students scoring at the meet and exceeds levels of the CRCT. Schools that had implemented whole-school reforms were compared to schools where no whole-school reform was implemented. The secondary purpose of the research was to determine the relative impact of each of the respective whole-school reform models on reading proficiency when compared to each other.

The analysis of the results is presented in three phases. Phase 1, addresses school level findings for Grades 3, 5 and 8 for African-American and disadvantaged students attending schools that implemented whole-school reading reform programs compared to students who attended schools that did not implement whole-school reading reform programs. The second phase examined the impact of the selected whole-school reforms on reading achievement. Finally, the analysis utilizes a pair-wise comparison between each reform to determine the most effective program. The study was guided by the research questions. The null hypotheses served the purpose of quantifying the responses to the research questions.

Summary and Discussion of School-Level RQ 1, 2, 3 and Ho 1-6.

Research Question 1. Is there a significant difference in the school level third grade Reading performance between African American or disadvantaged students in the Intervention

Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African American or disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

H_{o1}: There is no significant difference in the school level third grade Reading performance between African American or students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African American students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

The origin of this research question, as indicated in the review of the related literature, precedes the findings of the *Nation at Risk Report* (1983) that provided a disturbing picture of the state of education in America and indeed goes back at least to the start of Federal programs such as *Head Start* and *Follow Through* (Egbert, 1981; McLaughlin, 1995). Each of these individuals identified the need for educational reforms designed specifically to improve the academic proficiency of underserved by implementing programs or instructional strategies to close the achievement gaps that existed between certain ethnic and social groups. The specific research intent of this question is to determine the degree to which educational reforms in general can make a difference in the reading proficiency of African-American students regardless of the specific whole-school reform or grade level. The achievement gap within and between African-American students, as indicated both from an historical and contemporary perspective, is well researched and documented (American Psychological Association, 2011; Baker, 2005; Griffin, Jayakumar, Jones, & Allen, 2010).

The response to RQ1 supports the findings of Evans (2005), Ferebee (2005), and Freeman (2005) when they concluded that little has changed achievement-wise relative to the

implementation of the various educational reforms required under NCLB.

The statistical analysis failed to reject the null hypothesis (**Ho1,2**) of no difference. The results of the Mann-Whitney *U* resulted in a median rank of 24.25 for reform compared to a Median Rank of 20.75 for no reform schools ($p=.365$), indicating that there was no significant difference between the reading proficiency of third grade African-American students in schools that implemented a whole-school reform and African-American students attending schools where no whole-school reform was implemented. A similar finding occurred for Ho2, which addressed the same prediction with the control variable for disadvantaged third grade students ($p = .155$), resulting once again in a failure to reject the hypothesis of no difference.

Summary and Discussion of RQ2 and Ho3 and 4.

Research Question 2. Is there a significant difference in the school level fifth grade Reading performance between African American or disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African American or disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho3: There is no significant difference in the school level fifth grade Reading performance between African American students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African American students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

Research Question 2 is identical to RQ1 with one exception—RQ2's focus is on fifth-grade students rather than third-grade students. However, the similarity ends there because the literature relative to what difference whole-school reforms compared to no reforms have on

African-American reading proficiency is consistent in reporting that there is no significant difference between the impact of reform and non-reform relative to African-American students' reading proficiency (NAEP, 2011). Again, the results of the Mann-Whitney U concluded there was no significant difference based upon a Median Rank of 19.91 and 20.12, respectively, for reform and non-reform ($p = .955$), resulting in a failure to reject the null hypothesis (**H03**). Similar to African-American students, the reading proficiency of students from disadvantaged was not significantly impacted based upon their attending a reform or non-reform school.

H04. There is no significant difference in the school level fifth grade Reading performance between disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

The analysis of H04 also resulted in a failure to reject the null hypothesis where the p value was .336. The results of H04 and 5 is conclusive in their findings relative to the impact of schools with whole-school reading reforms and schools where no whole-school reform has been implemented. The reader must exercise some caution, however in the interpretation due to the fact that the independent variables were not specifically identified or taken into account during the analysis. Still, the basic research question of whether whole-school reform in general impacts the reading proficiency of African-American and or disadvantaged eight grade students.

Summary and Discussion for RQ3 and H05 and 6.

Research Question 3. Is there a significant difference in the school level eight grade Reading performance between African American or disadvantaged students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core

Knowledge is implemented and African American or disadvantaged students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

Ho5: There is no significant difference in the school level eight grade Reading performance between African American students in the Intervention Group where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented and African American students in the Comparison Group where Success for All, Direct Instruction, or Core Knowledge is not implemented.

The Mann-Whitney U for eighth grade African-American students in reform and non-reform schools revealed that there is no statistical significant difference in reading proficiency relative to grade level. The analysis resulted in a Median Rank of 10.72 for non-reform performance compared to a rank of 8.28 for schools with reforms ($p = .329$), resulting in a failure to reject the null hypotheses. Similar to the results of the analysis of African-American and disadvantaged students in the third and fifth grades, the reading proficiency for grade 8 students did not vary based upon the presence or absence of whole-school reading reform programs. Again, the small number of schools must be taken into account when interpreting the results.

Student-Level Summary and Discussion of RQ 4, 5, 6 and Ho 7-12.

Research Questions 4, 5, and 6, as did Research Questions 1, 2 and 3 speaks to the basic question of whether or not whole-school reading reforms in general make a difference in reading proficiency. The major differences between the two sets of research questions pertain to race and socioeconomic status. In effect, does the impact of reform versus non-reform schools differ with African-American or disadvantaged students? Another difference between the two sets of research questions is the fact that Questions 1, 2, and 3 represent school-level data and analysis, while the latter set—Questions 4, 5, and 6—represent student-level data and analysis.

The researcher utilized a chi-square test in order to determine whether there were significant differences in proficiency rates between African-American and disadvantaged students in reform and non-reformed schools. The chi-square test examines the prediction of the Null Hypothesis that the proportion of students attaining proficiency is equal for students in reformed and non-reformed schools. The operative word or control variable for RQ 4, 5, and 6 is disadvantaged students.

Hypotheses 7 and 8 examined the differences in the reading proficiency of African-American and disadvantaged third grade students in reform and non-reform schools using student level data. The results of the chi-square statistic of 4.461 ($p=.035$) for African-American students and 4.833 ($p=.028$) for disadvantaged students respectively resulted in a rejection of the null hypothesis of no difference. It should be noted that the large sample sizes may have influenced the significant results. On the other hand, the large sample size generally reduces the chances for committing a Type II error. The small effect size of .04 and .05 respectively, requires caution in the interpretation of the results.

Hypotheses 9, 10 used student data to the impact of reform and non-reform programs on African-American and disadvantaged students in grade . There were no significant differences found for either African-American or disadvantaged students relative to reading proficiency.

Grade 8 results, Research Question 6, Ho 11 and 12, produced mixed results. African-American student's reading performance differences between reform and non-reform students were found not to be significant ($p=.55$). The results for disadvantaged students attending schools with whole-school reading reform programs resulted in significant differences where chi-square was 8.347 and $p<.04$. In effect, disadvantaged students in grade 8 attending schools with whole-school reading programs did significantly better than disadvantaged students

attending schools without whole-school reading reform programs. An effect size of only .13 however suggest that the finding may have been caused by other factors.

Connection of the Findings to the Literature

The literature is replete with studies that chronicle the impact of poverty or being disadvantaged on academic achievement (Strayhorn, 2010; Yeung & Conley, 2008). The majority of the literature concludes that while there is a general consensus that being disadvantaged is often an impediment to achievement, whole-school reforms can have positive results (Daugherty, 2011; Viadero, 2010). Their findings (Strayhorn, 2010; Yeung & Conley, 2008) to some extent, support findings for RQ4; namely, proficiency rates were significantly higher (chi-square $d = 4.833, p = .028$) for disadvantaged students attending schools with whole-school reforms compared to schools without reforms resulted in the rejection of *Ho4* indicating that there is a significant difference in reading proficiency when disadvantaged students attend reformed schools. According to some researchers, this kind of result is more likely to occur when the whole-school reforms are based on political factors rather than pedagogy (Fullan, 2010; Grindle, 2004; Neague, 2011). The assertion that politics sometimes drives the respective educational reforms implemented is supported by (Carnine, 1983) when comments that “so-called experts often constrain objective research (p. 2). There were, however, different conclusions for disadvantaged students in Grade 5 where the chi-square ($d = 2.832, p = .092$) resulted in a failure to reject the hypothesis of no difference (**Ho5**). Some of the research findings reported in the literature review (Thompson & Thompson, 2000) suggested that the reasons contributing to the mixed results between reform and non-reform initiatives maybe caused by a lack of fidelity in the implementation process or faulty research designs.

Research Question 6. Is there a significant difference in the student level eight grade reading performance between African-American or disadvantaged students in the intervention group—where the whole-school reform of Success for All, Direct Instruction, and Core Knowledge is implemented—and African-American or disadvantaged students in the comparison group where Success for All, Direct Instruction, or Core Knowledge is not implemented?

The chi-squared statistic for disadvantaged students was 8.347, $p=.04$, resulting in an acceptance of the null hypothesis. Like the results from Grade 3, disadvantaged students did better in schools implementing whole-school reforms than they did in schools with no reforms. In addition, the Phi coefficients (effect size) was .06, suggesting that there is a weak relationship between the impacts of reform versus non-reform schools for eighth grade disadvantaged students.

Summary and Discussion of RQ7

Research Question 7. Is there a significant difference between the relative impact of Success for All, Direct Instruction, or Core Knowledge on the reading proficiency of African-American or disadvantaged students following a pairwise comparison between each of the reform models?

Research Question 7 represents an important component of the overall purpose of this research study. It specifically addresses, from a practical perspective, what educational leaders and practitioners really want and need to know relative to helping students achieve; namely, what works? The question provides quantitative data to justify adopting or not adopting various whole-school reading reform programs for underserved students such as African-American and disadvantaged students. This approach is supported by Marzano, Pickering, and Pollock (2001),

who concluded that, “although we should draw no hard and fast conclusions...it illustrates the need to study the effects of instructional strategies on specific types of students in specific situations, with specific subject matter” (p. 9). Clewell et al. (2007), Harford (2007), and Rothstein (2008) augmented this assessment by asserting that it stands to reason that if causes and reasons differ for the achievement of students, so should the strategies and methodologies.

Summary and Discussion of the comparison of Reform Models

The reformed schools in the sample followed one of three models: Success for All (SFA), Direct Instruction (DI), and Core Knowledge (CK). The next stage of the analysis compared proficiency rates between each of the three models at the school level.

Direct Instruction (DI) versus Success for All (SFA). As shown in Table 9, the percentage of students attaining proficiency was significantly higher ($U = 33.50, p < .001$) in schools implementing the Direct Instruction reform model ($M = 88.34$) compared with those that were implementing the Success for All reform model ($M = 81.81$). The probability that a randomly drawn score from the Direct Instruction population will be greater than one drawn from the Success for All population is .13, resulting in the rejection of the H_0 . There is currently a host of evaluation studies indicating the effectiveness of DI with various populations (Dale, Philip, & Cole, 1988; Lewis, 1982; Rawl, Ruth, & France, 1982). SFA has also conducted an impressive string of program evaluations in support of the effectiveness of SFA. One major difference between the two whole-school reform programs is that unlike DI, SFA’s evaluations targeted various ethnicities and at risk students (Slavin & Madden, 1995; Venezky, 1994).

Core Knowledge (CK) versus Success for All (SFA). As shown in Table 8, the percentage of students attaining proficiency was significantly higher ($U = 11.00, p < .001$) in schools implementing the Core Knowledge reform model ($M = 91.44$) compared with the

Success for All reform model ($M = 81.81$). The probability that a randomly drawn score from the Core Knowledge population will be greater than one drawn from the Success for All population is .13 resulting in a rejection of H_0 , there is a significant difference between the two reform models when compared to each other.

Core Knowledge (CK) versus Direct Instruction (DI). Table 9 illustrates the differences in proficiency rates between schools implementing Core Knowledge and Direct Instruction reform models. The probability that a randomly drawn score from the Direct Instruction population will be greater than one drawn from the Success for All population is .04. Proficiency rates differed significantly between reform models. Significantly higher proficiency rates were found in schools implementing the Direct Instruction or the Core Knowledge reform models compared with the Success for All model indicating that the null hypothesis of no difference is not rejected. The literature pertaining to both Core Knowledge and Direct Instruction is somewhat mixed (Adams et al., 2004; Lindsey, 2010). Basically, the CK curriculum is based on the belief that systematic phonics-based instruction involving decoding skills and read-aloud exercises build oral language (Bradley, 2005; Hirsch, 2008). Direct Instruction, on the other hand, includes explicit, systematic instruction based on scripted lesson plans (Adams & Engelmann, 2007; Engelmann, 1996; Maccoby & Zellner, 1970). There appears to be some evidence that DI is more appropriate for disadvantage and at-risk students and some of their evaluations appear to support this contention (Dale, Philip, & Cole, 1988).

Limitations of the Study

The major limitation of this study was the inability to control for the fact that there could be an interaction between the control variables of African-American and disadvantaged students. The interaction being referred to is the fact that it is possible that the African-American students

in the sample population may also be disadvantaged. In order to remove or reduce the limitations, African-American student's socioeconomic status would have to be isolated. The State Department of Education provides the student data by race but does not disaggregate race by economic status. A related limitation and concern are the numerous variations relative to the meaning and understanding of what constitutes being *economically disadvantaged* (Census Bureau, 2009; Clewell et al., 2007; Rist, 1973). Albeit the State Department utilized a standard set of criteria to categorize the socioeconomic status of students, there remains a great many grey areas (Kulm, 2007; Strayhorn, 2010; Yeung & Conley, 2008). The grey areas pertaining to the categorization of being disadvantaged include, but are not limited to cultural, educational attainment, family structure and ethnicity (Bensalem, 2011; Cooper, Crosnoe, Suizzo, & Pituch, 2010). For example, students from urban areas are often considered to be culturally disadvantaged in spite of the fact that many urban communities include individuals and families from various economic groups. A similar situation exists for individuals who have limited educational backgrounds. In some situations, students from single-family homes are considered to be disadvantaged regardless of their economic status. For purposes of clarity, this study utilized the Georgia State Department of Education's criteria and definition for disadvantaged as only referring to *economic disadvantaged as opposed to being physically or mentally disadvantaged*. The Georgia definition and categorization comes directly from the Federal Government.

A second limitation of this study involves the assumption of fidelity relative to the implementation of the respective whole-school reform models. In effect, the researcher can only rely on the school principal's or training director's assurance that the program was implemented as prescribed by the creator of the program (Kovaleski, Gickling, & Marrow, 1999).

While it may appear that in view of the fact that each of the reform programs examined in the study are different in very specific ways regarding scope and intent, in truth, program differences may also be a strength as well as a limitation. For example, both Core Knowledge and Direct Instruction supervise their instructional interventions, and sequence their program components, Direct Instruction pays particular attention to the implementation phase of their program (Engelmann, 2007; Feinzimer, 2011). Core Knowledge, on the other hand, places more emphasis on the sequence of the pedagogy (Hirsch, 2008).

While the sample size for the student analysis was large and sufficient, the number of schools actually implementing each of the respective reforms was relatively small. In view of the fact that a small sample could impact the statistical power, thereby increasing the chance of committing a Type I or Type II error and incorrectly assuming that the obtained results are representative of the larger population (Gall, et al., 2003). In effect, the researcher's intent is to increase the level of confidence, both statistically and practically. The researcher addressed the statistical power limitation by conducting pre-assessment of sample size necessary to obtain the level of confidence desired and concluded that the sample size overall was reasonable (Gay, Mills, & Mills, 2009; Thompson, 2007). Another possible limitation is the fact that there is no pre-measure or assessment of students' achievement. Again, the researcher attempted to offset this limitation by matching schools relative to both achievement and demographics. Both the reliability and validity of the CRCT were in an acceptable range with the total test reliability for 2011 ranged from 0.89 to 0.90 for Reading, 0.85 to 0.91 for English/Language Arts, and 0.87 to 0.91 for Mathematics. The GA DOE is confident that the CRCT are both reliable and valid (Georgia Department of Education, 2011).

Implications

In view of the fact that the issue of the state of education in general and especially in underserved groups such as minorities, disadvantaged students, and students with disabilities, the results of this study have several clear and important implications for the educational community. The three major questions addressed in the study relative to what degree reforms impact reading achievement among African-American and disadvantaged students could have major implications for how we close the achievement gaps. Specifically, while it is anticipated that the *No Child Left Behind Act* (2001) may be changed or discarded altogether, (U.S. Department of Education, 2010), to date, the mandate to better serve the underserved students, remains in place. Further, the requirement that school districts only adopt programs that are scientifically research-based programs has a direct connection to this study's findings relative to the quality of the three reform models examined. In addition, the findings could have direct impact on curriculum and instruction. Some of the implications from this study can be summed up in a study by Marzano, Pickering, and Pollock (2001) who stated the following:

Although our synthesis of the research has taught us a great deal, there are still many questions as yet unanswered. Some are:

Are some instructional strategies more effective in certain subject areas?

Are some instructional strategies more effective with students from different backgrounds?

Are some instructional strategies more effective with students of different aptitude? (p. 8)

This study addresses most of Marzano et al.'s (2001) questions and even provides insight into questions not posed. For example, do whole-school reforms differentiate by grade level?

Recommendations for Future Research

The most obvious and logical place to start in making recommendations for future research is with the identified limitations of the study. Therefore, any future research addressing the scope and purposes of this study might include a more complete disaggregation of the subgroups by identifying a third subgroup consisting of African-American students who are also disadvantaged. In effect, there would be three control variables:

- African-American students who are economically disadvantaged
- African-American students who are not economically disadvantaged
- Disadvantaged students

The rationale here is to determine the impact of race and poverty. Another recommendation for future research would be to control for the diversity in the respective reform models by comparing like-models with each other and with models that are most dissimilar. The discovery process might also include comparing the impact of various reforms when gender is a control variable (Blum, 1997; Rush, 2005). Future research could be more value-added by examining the performance of African-American and disadvantaged students across multiple years. This longitudinal approach would help to make the results more reliable, thereby improving the internal validity by controlling for the impact of extraneous variables (Gall, Gall, & Borg, 2003; Huck, 2012; Ross et al., 2004).

While a study that could control for the fidelity of the implementation process of reform models would be desirable, the human and financial costs may be prohibitive. Finally, the examination of multiple school years could add a degree of reliability to the results and might even provide some formative aspects to improving the reading proficiency of African-American and disadvantaged students.

Conclusions

The goal of the study was to address long-standing problems associated with the improvement of reading proficiency among African-American and disadvantaged students by examining the impact of selected whole-school reading reform programs. African-American and disadvantaged students were selected as the control variables for several reasons. Firstly, the aforementioned groups have historically been among the lowest relative to academic achievement in most subject areas, not just reading (Clewell et al., 2007; Rist, 1973). Secondly, NCLB (2001) mandates that these particular groups have been and continues to be underserved, and in many instances, *they have been left behind* (Berends et al., 2002; Williams, 2006). The general consensus is that American students are still lagging behind students from other industrialized countries and the gap does not appear to be closing significantly. In short, the evidence appears to conclude that researchers and educators do not yet have all of the answers for closing the achievement gaps between various races, ethnicities, socioeconomic groups or genders (U.S. Department of Education, 2006; Wenglinsky, 2004). Along similar lines, although the research does provide evidence that there are effective instructional reading strategies and programs such as DI, CK and SFA, there is not much evidence to indicate there is any single approach that works in all situations (Bensalem, 2011; Clewell, Campbel & Perlman, 2007; Delpit, 2007). In spite of the the issues and problems previously identified, some progress is being made. This is due in part to the fact that educational improvement is both a national and a local priority. The progress is also helped along by the involvement of more classroom teachers and other support staffs. And to no small measure, universities and colleges are encouraging and promoting more *grounded* research designs and more *action* research.

It is fair to conclude, however, that there is another school of thought that advances the belief that current education reform research is flawed because it is dominated by too much descriptive and qualitative research (Carnine, 1983, 1984; Engelmann, 2007). Professor Carnine, concludes that educational research should look more to empirical, rigorous and scientific approach to evaluating the effectiveness of education reform strategies and programs. The findings from this study, which indicates that Direct Instruction is among the most rigorously evaluated education reforms program in existence (Education Consumers Foundation, 2011; What Works Clearinghouse, 2007). This assertion is supported by a host of evidence over a 25 year period (Adams & Engelmann, 2007; Borman, & Hewes, 2002). Further, the research on Direct Instruction has included effect sizes of its findings, thereby adding statistical evidence to support the strength of the relationship between the outcomes of Direct Instruction and achievement. An effect size of .25 is typically considered to represent a significant relationship between variables (Education Consumers Foundation, 2011; Hattie, 2009). A meta-analysis study on Direct Instruction conducted by White, (1988), found an average size of .84 on the effect of Direct Instruction on special education students. Finally, based upon the evidence provided by the literature pertaining to the effectiveness of education reforms and the findings from this study, it is concluded that overall, Direct Instruction provides educators with the greatest opportunity to advance and improve the education of the largest number of students.

The results of this study do provide evidence to warrant further study relative to the impact of whole-school reform on reading proficiency. As indicated in more than one instance, school implementing reforms do have a greater impact on both African-American and disadvantaged students, regardless of which of the three reform models examined. The findings were also consistent in ranking the reforms based upon their overall effectiveness. In fact, the

vast majority of the reform evaluations reviewed indicated that Direct Instruction tended to produce great reading gains for low income students and students with disabilities than did either Success for All or Core Knowledge (Adams & Engelmann, 1996; Bradley, 2005; Englemann, 1968; Huitt, 1996; Ryder, Burton, & Silberg, 2006). The research evidence provided above regarding the effectiveness of Direct Instruction with African-American and disadvantaged students is consistent with many of the findings discussed in this study (Goldman, 2000; Watkins, 2008).

REFERENCES

- Adams, G. L., & Engelmann, S. (2007). *Research on direct instruction: 25 years beyond DISTAR*. Seattle, WA: Educational Achievement Systems.
- Adams, G. L., Slocum, T., et al. (2004). A critical review of Randall Ryder's report of direct instruction reading in two Wisconsin school districts. *Journal of Direct Instruction*, 4(2), 161-190.
- Adams, J. (2005). No Child Left Behind: An investigation of the perceptions of public school administrators responding to the impact of United States legislation on their schools. *Dissertation Abstracts International-A*, 66, 05.
- Aguinis, H., & Pierce, C. A. (2006). Computation of effect size for moderating effects of categorical variables in multiple regression. *Applied Psychological Measurement*, 30(5), 440-442.
- Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, 8(2), 606-621.
- American Psychological Association. (2011). *Fact sheet: Education and socioeconomic status*. Retrieved from <http://www.apa.org/pi/ses/resources/publications/factsheet-education.aspx>.
- American Psychological Association. (2009). *Publication manual* 6th ed Edition. Washington, DC.
- Andrews, T. (n.d.). *Educational reforms: Core knowledge*. Retrieved from <http://www.articlesnatch.com/Article/Educational-Reforms--Core-Knowledge/924952>.
- Anyon, J., & Greene, K. (2007). No Child Left Behind as an anti-poverty measure. *Teacher Education Quarterly*, 34(2), 157-62.

- Archibald, S. (2006). Narrowing in on educational resources that do affect student achievement. *Peabody Journal of Education*, 81(4), 23-42.
- Armstrong, A. (2010). Myths of poverty. *The education digest*, 75(8), 49-53.
- Avila, C. (2010). *Examining the implementation of district reforms through gap analysis: Addressing the performance gap at two high schools*. (Doctoral dissertation). Retrieved from Dissertations & Theses: Full Text. (Publication No. AAT 3465946).
- Bahr, P. R. (2010). Preparing the underprepared: An analysis of racial disparities in postsecondary mathematics remediation. *Journal of Higher Education*, 81(3), 209-237.
- Baker, J. (2009). Class, ability, mobility: Economic and academic paths from middle school to early adulthood. *Journal of Education Finance*, 34, 355-371.
- Baker, J. D., Rieg, S. A., & Clendaniel, T. (2006). An investigation of an after school math tutoring program: University tutors + elementary students = a successful partnership. *Education*, 127(5), 287-293.
- Baker, P. (2005). The impact of cultural biases on African-American students' education: A review of research literature regarding race-based schooling. *Education and Urban Society*, 37(3), 243-256.
- Becker, W.C., & Engelmann, S. (1996). Sponsor findings from Project Follow Through. *Effective School Practices*, 15(4), 33-42.
- Beghetto, R. (2003) *Scientifically Based Research*. ERIC Clearinghouse on Educational Management. Retrieved February 20, 2013.
- Bensalem, M. D. (2011, September 13). A nation at risk: A failure to reconnect with founders' principles. *Phillyburbs.com*. Retrieved from: http://www.phillyburbs.com/news/local/courier_times_news/opinion/guest/a-nation-at-

risk-a-failure-to-reconnect-with-founders/article_42e6d13f-3dbc-5893-9a3c-6ca2f41c45a8.html

Bereiter, C. & Kurland, M. (1996). A constructive look at Follow Through results. *Effective School Practices, 15*(1), 1.

Berends, M., Bodilly, S., & Kirby, S. N. (2002). *Facing the challenges of whole-school reform: New American schools after a decade*. Santa Monica, CA: RAND.

Berlak, H. (2001). *Race and the achievement gap*. Retrieved from <http://www.rethinkingschoolsonline.org>

Berliner, D. C. (2006). Our impoverished view of educational research. *Teachers College Record, 108*(4), 949-95.

Birenbaum, M., Nasser, F., & Tatsuoka, C. (2007). Effects of gender and ethnicity on fourth graders' knowledge states in mathematics. *International Journal of Mathematical Education in Science and Technology, 38*, 301-319. doi: 10.1080/00207390601116052

Blum, D. (1997). *Sex on the Brain: The biological differences between men and women*. New York: Viking

Bock, G., Stebbins, L., & Proper, E. (1977). *Education as experimentation: A planned variation model (Volume IV-A & B) Effects of follow through models*. Washington, DC: ABT Associates.

Bond, L. A. (1996). Norm- and criterion-referenced testing. *Practical Assessment, Research & Evaluation, 5*(2).120-125. Retrieved from <http://PAREonline.net/getvn.asp?v=5&n=2>

Borman, G. D. et al. (2000). *Four models of school improvement successes and challenges in reforming low-performing, high-poverty Title I schools*. Retrieved from www.csos.jhu.edu.

- Borman, G.D., & Hewes, G.M. (2002). The long-term effects and cost-effectiveness of success for all. *Educational Evaluation and Policy Analysis*, 24,(3), 243-266. Retrieved from http://www.successforall.net/_images/pdfs/Third_Year_Results_06.doc.
- Borman, G. D., Slavin, R. E., Cheung, Alan C.K.,Chamberlain, A.M., Madden, N.A., Chambers, B. (2006). Final reading outcomes of the national randomized field trial of Success for All. *Educational Research Journal*, 4 (492), 701-731.
- Borman, G.D., Slavin, R.E., Cheung, Alan C.K.,Chamberlain, A.M., Madden, N.A., Chambers, B. (2010). The impact of No Child Left Behind on students, teachers, and schools. *Brookings*. Retrieved from <http://www.brookings.edu>.
- Bowen, G., & Bok, D. (1998). *Shape of the river*. Princeton, NJ, Princeton University Press.
- Bradley, R. H. (2005). Evaluation of the core knowledge program in Arkansas. *Core Knowledge*. Retrieved from: http://coreknowledge.org/mimik/mimi_uploads/documents/25/CK_EvalRpt_AR2005.pdf
- Braun, H., Chapman, L., & Vezzu, S. (2010). The black-white achievement gap revisited. *Education Policy Analysis Archives*, 18(21), Retrieved from <http://epaa.asu.edu/ojs/article/view/772>
- Burchinal, M., McCartney, K., et al. (2011). Examining the black-white achievement gap among low-income children using the NICHD study of early child care and youth development. *Child Development*, 82(1), 1404-1420.
- Burney, V. H., & Beilke, J. R. (2008) .The constraints of poverty on high achievement. *Journal for the Education of the Gifted*, 31(3), 171-197.
- Burton, L. (2007). Childhood adultification in economically disadvantaged families: A conceptual model. *Family Relations*, 56(2), 329-345.

- Caldas, S. J., Bernier, S., & Marceau, R. (2009). Explanatory factors of the black achievement gap in Montreal's public and private schools: A multivariate analysis. *Education and Urban Society, 41*(2), 197-215.
- Carlyle, E. A. (2008). *Achievement disparity for minority and economically disadvantaged student populations: A limited analysis of No Child Left Behind*. (Unpublished doctoral dissertation).Walden University, Minneapolis, MN.
- Carnine, D. W., Silbert, J., &Kameenui, E. J. (1997). *Direct instruction reading* (3rd ed.). Upper Saddle River, NJ: Merrill/Prentice-Hall.
- Carpenter II, D. M., Ramirez, A., & Severn, L. (2006). "Gap" or "gaps": Challenging the singular definition of the achievement gap. *Education and Urban Society, 39*(1), 113-127.
- Carroll, R.S., & Carroll, J.D. (2002). *Statistics made simple for school leaders: Data-driven decision making*. Rowman & Littlefield Publishers, Inc.
- Center on Educational Policy. (2012). Retrieved from <http://www.cep-dc.org/>
- Center for Educational Statistics, U.S. Department of Education. (2011). *Reading 2011 National assessment of educational progress at grades 4 and 8*. Retrieved from <http://nces.ed.gov/nationsreportcard/pubs/main2011/2012457.asp>
- Center for Research and Reform in Education. (2005). *Randomized evaluation of success for all*. Retrieved from <http://education.jhu.edu/partnerships/crre.html#evaluation>
- Chapman, L. (2007). An update on No Child Left Behind and national trends in education. *Arts Education Policy Review, 109*(1), 25-36.
- Charles, C. Z., Fischer, M. J., Mooney, M. A., & Massey, D. S. (2009). Affirmative-action programs for minority students: Right in theory, wrong in practice. *The Chronicle of*

- Higher Education*, 55(29), A29-A30.
- Chaturvedi, M. (2009). School environment, achievement motivation, academic achievement. *Indian Journal of Social Science Research*, 6(2), 29-37.
- Church, M. B. (2000). Comprehensive school reform demonstration program: The third wave and its preliminary effects. *Pacific Resources for Education and Learning*. Retrieved from <http://www.gpoaccess.gov/eric/200212/ed465199.pdf>.
- Clayton, J. K. (2011). Changing diversity in U.S. Schools: The impact on elementary student performance and achievement. *Education and Urban Society*, 43, 671-695.
- Clewell, B. C., Campbell, P. B., & Perlman, L. (2007). *Good schools in poor neighborhoods: Defying demographics, achieving success*. Washington, DC: The Urban Institute Press.
- Coe, R. (2002). *It's the effect size, stupid: What effect size is and why it is important*. Paper presented at the Annual Conference of the British Educational Research Association. Retrieved from www.leeds.ac.uk/educol/documents/00002182.htm.
- Cohen, G. (2009). Study: Writing about values boosts grades, shrinks achievement gap. *PBS Newshour*. Retrieved from: http://www.pbs.org/newshour/updates/science/jan-june09/achievement_04-17.html
- Cohen, M. (2003). *Nation at risk twenty years later*. Retrieved from [http://www.aei.org/EM/Static/Static Page/260?page=Summary](http://www.aei.org/EM/Static/Static%20Page/260?page=Summary)
- Coladarci, T., Cobb, C.D., Minium, E.W., Clarke, R.B. (2011). *Fundamentals of Statistical Reasoning in Education* (3rd ed.). Danvers, MA.
- College Board. (2007). *Recommendations for reauthorization of the Elementary Secondary Education Act*. New York, NY: The College Board.

- Cooper, C. E., Crosnoe, R., Suizzo, M. A., & Pituch, K. A. (2010). Poverty, race, and parental involvement during the transition to elementary school. *Journal of Family Issues, 31*(3), 859-883
- Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Belmont, CA: Wadsworth.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.
- Cruz, P. (2011). *Implementing a mandated school reform model in low-performing schools: A multiple-case study of four New Mexico schools and their leaders*. (Doctoral dissertation). Retrieved from Dissertations & Theses: Full Text. (Publication No. AAT 3296121)
- Cutler, P. (1984), The measurement of poverty: A review of attempts to quantify the poor with special reference to India. *World Development, 12*(11/12), 1119-1130.
- Dale, P. S., and Kevin, N. C. (1988). Comparison of academic and cognitive programs for young handicapped children. *Exceptional Children, 54*(5), 439-447.
- Daniels, C., Madden, N.A., & Slavin, R.E. (2004). *The success for all middle school: Third year evaluation report*. Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from <http://www.eric.ed.gov/PDFS/ED483813.pdf>
- Daugherty, A. (2011). *A study of the learning-focused school improvement model and its*

- effects on third grade reading scores in a suburban, metropolitan school system.*
Educational Policy Studies Dissertations. Paper 77. Retrieved from
http://digitalarchive.gsu.edu/eps_diss/77
- Dee, T., & Jacob, B. (2010). *The impact of No Child Left Behind on students, teachers, and schools.* Washington, DC: National Bureau of Economic Research.
- Delpit, L. (2007, May 21). Evaluating No Child Left Behind. *The Nation*. Retrieved from
<http://www.thenation.com/doc/20070521/darling-hammond>
- Desimone, L. M., Smith, T. M., & Frisvold, D. (2007). Teacher quality, evidence-based practice, and students in poverty. In A. Gamoran (Ed.), *Standards-Based Reform and the Poverty Gap: Lessons for "No Child Left Behind"* (p. 2-6). Washington, DC: Brookings Institution Press.
- Douglas, A. D. (2011). *A Study of the Learning-Focused Model and its effectiveness on third grade reading scores in a suburban, metropolitan school system.* (Unpublished doctoral dissertation). Georgia State University, Athens.
- Drukker, M., Feron, F. J. M., Mengelers, R., & Van Os, J. (2009). Neighborhood socioeconomic and social factors and school achievement in boys and girls. *Journal of Early Adolescence, 29*(1), 285-306.
- Duncan, G. J., & Magnuson, K. A. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *Future of Children, 15*(1), 35-54.
- Dunn, J. L., & Allen, J. (2009). Holding schools accountable for the growth of non-proficient students: Coordinating measurement and accountability. *Educational Measurement: Issues and Practice, 28*(4), 27-41.
- Easton, V.J., & McColl, J.H. (n.d.). *Statistics Glossary*. Retrieved from

http://www.stats.gla.ac.uk/steps/glossary/confidence_intervals.html#confinterval

Eddy, C. M., & Easton-Brooks, D. (2011). Ethnic matching, school placement, and mathematics achievement of African-American students from kindergarten through fifth grade. *Urban Education*.46 (6), 1280-1299.

Educational Research Services. (2001). *What can schools do to reduce the achievement gap?*
Retrieved from: <http://www.ers.org/otsp/otsp3.htm>.

Egbert, R.L. (1981). *Some thoughts about Follow Through thirteen years later*. Lincoln: Nebraska University. (ERIC Document Reproduction Service No. ED244733). Retrieved from http://en.wikipedia.org/wiki/Project_Follow_Through

Encyclopedia of American Education. (2011). A nation at risk: The imperative for educational reform. *American Education*. Retrieved from: http://american-education.org/1381_a_nation_at-risk_the-imperative-for-educational-reform.html.

Engelmann, S. (2007). *Teaching needy kids in our backward system: 42 Years of Trying*. Eugene, Oregon: ADI Press.

Engelmann, S. (1996). Relating operant techniques to programming and teaching. *Journal of School Psychology*, (1), 89-96

Evans, R. (2005). Reframing the achievement gap. *Phi Delta Kappan*, 86, 590-593.

Feinzimer, L. (2011). *National comprehensive school reform: An analysis of six reform models 1980—2000*. (Doctoral dissertation). Retrieved from Dissertations & Theses: Full Text. (Publication No. AAT 3367105).

Fidler, F. & Thompson B. (2001). Computing correct confidence intervals for ANOVA fixed and random-effects effect sizes. *Educational and Psychological Measurement*, 61(1), 575-604

- Fischer, G. (2010). *School reform for 21st-century skills under the No Child Left Behind accountability system: Perceptions of superintendents*. (Unpublished doctoral dissertation). Fordham University, New York.
- Forsee, B. (2010). Comprehensive school reform implementation: A gap analysis inquiry project for Rowland Unified School District. (Doctoral dissertation). Retrieved from Dissertations & Theses: Full Text. (Publication No. AAT 3465993).
- Frankenberg, E. (2009). The demographic context of urban schools and districts. *Equity & Excellence in Education, 42*(2), 255-271.
- Freeman, E. (2005). No Child Left Behind and the denigration of race. *Equity & Excellence in Education, 38*(2), 190-199.
- Fullan, M. (2010). *All systems go: The change imperative for whole system reform*. Thousand Oaks, CA: Corwin Press.
- Gall, M.D., Gall, J. P., & Borg, W. R. (2004). *Educational research: An introduction*. New York, NY: Longman.
- Gardner, W. (2012). Income more important than race in achievement gap. *Education Week*. Retrieved from:
http://blogs.edweek.org/edweek/walt_gardners_reality_check/2012/02/income_more_important_than_race_in_achievement_gap.html?cmp=SOC-SHR-FB
- Gay, L. R., Mills, G. E., & Mills, P. (2009). *Educational research: Competencies for analysis and applications*. Upper Saddle River, NJ: Pearson.
- Geoffroy, M.C. et al. (2010). Closing the gap in academic readiness and achievement: The role of early childcare. *Journal of Child Psychology and Psychiatry, 51*, 1359-1367.
- Georgia Department of Education. (2007). *Criterion-Referenced Competency Tests*. Retrieved

- from: http://www.doe.k12.ga.us/ci_testing.aspx?PageReq=CI_TESTING_CRCT
- Georgia Department of Education. (2010). *Criterion-Referenced Competency Tests*. Retrieved from: http://www.doe.k12.ga.us/ci_testing.aspx?PageReq=CI_TESTING_CRCT
- Georgia Department of Education. (2011). Validity and reliability for the 2011 Criterion-Referenced Competency Tests. An Assessment & Accountability Brief: 2011 CRCT Validity and Reliability.
- Georgia Professional Standards Commission (2009). *An overview of Georgia educator workforce*. Retrieved from: <http://www.gapsc.com/Research/2008Report/AnOverviewOfGeorgia%27sEducator>.
- Gersten, R., Beckmann, S., et al. (2009). *Assisting students struggling with mathematics: Response to intervention for elementary and middle schools*. Retrieved from <http://ies.ed.gov/ncee/wwc/practiceguide.aspx?sid=2>
- Glass, G.V., & Camilli, G. A. (1981). *"Follow through" evaluation*. Washington, DC: National Institute of Education.
- Goddard, R. D., Salloum, S. J., & Berebitsky, D. (2009). Trust as a mediator of the relationships between poverty, racial composition, and academic achievement: Evidence from Michigan's public elementary schools. *Educational Administration Quarterly*, 45(2), 292-311.
- Griffin, K. A., Jayakumar, U. M., Jones, M. M., & Allen, W. R. (2010). Ebony in the ivory tower: Examining trends in the socioeconomic status, achievement, and self-concept of black, male freshmen. *Equity & Excellence in Education*, 43(2), 232-248.
- Grindle, M. (2004). *Despite the odds: The contentious politics of education reform*.

- Princeton, NJ: Princeton University Press.
- Grissom, R.J., & Kim, J. J. (2012). *Effect sizes for research: Univariate and multivariate applications* (2nd ed.). New York: Routledge.
- Hagenaars, A., and de Vos, K. (1988). The definition and measurement of poverty. *The Journal of Human Resources*, 23(2), 212-221.
- Harford, T. (2007). The achievement gap. *Forbes.com*. Retrieved from http://www.forbes.com/2007/02/25economics-discrimination-education-lead_achieve07_cz_th_0301discrimination.html.
- Harris, N. D., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics*, 95, 789-798.
- Heritage Foundation. (2010). Retrieved from <http://www.heritage.org/>
- Hernstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York, NY: The Free Press.
- Hess, F. M. (2004). *Common sense school reform*. New York, NY: Palgrave MacMillan.
- Hess, F. M. (1999). *Spinning wheels: The politics of urban school reform*. Washington, DC: Brookings Institute.
- Hirsch, E. D. (2008). Annual Report 2008: The Core Knowledge Foundation. *Core Knowledge*. Retrieved from: <http://www.coreknowledge.org/ed-hirsch-jr>
- Hoff, D. (2000, September 6). Gap widens between black and white students on NAEP. *Education Week*. Retrieved from www.edweek.org.
- House, E., Glass, G., McLean, L., & Walker, D. (1978). No simple answer: Critique of the FT evaluation. *Harvard Educational Review*, 48(2), 128-160.

- Retrieved from: <http://pages.uoregon.edu/adiep/ft/grossen.htm>
- Huck, S. C. (2012). *Reading statistics and research*. New York, NY: Pearson.
- Hughes, C., Manns, N. N., & Ford, D. Y. (2009). Racial identity attitudes and academic achievement among at-risk black female adolescents. *The Journal of At-Risk Issues, 15*(1), 25-32.
- Huitt, W. (1996). Measurement and evaluation: Criterion- versus norm-referenced testing. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. Retrieved from: <http://www.edpsycinteractive.org/topics/measeval/crnmref.html>
- Huitt, W. (2008). *Direct instruction: A transactional model*. Valdosta, GA: Valdosta State University. Retrieved from <http://www.edpsycinteractive.org/topics/instruct/instevnt.html>
- Institute of Education Sciences. (2007). WWC intervention report Success for All July 2, 2007 What Works Clearinghouse English Language Learners. Retrieved from: http://ies.ed.gov/search/?output=xml_no_dtd&client=ies&site=ies&q=success+for+all
- Jackson, W. (2006). *An investigation of the impact of the Success for All whole-school reform model on the Elementary School Proficiency Assessment and the New Jersey Assessment of Skills and Knowledge in an urban district*. (Doctoral dissertation). Retrieved from Dissertations & Theses: Full Text. (Publication No. AAT 3301057).
- Jones, P. (2011). *The relationship between socialized mathematics instruction and academic performance in adolescent African-American male students*. (Doctoral dissertation). Retrieved from Dissertations & Theses: Full Text. (Publication No. AAT 3445105).
- Kim, J. S., & Sunderman, G. L. (2005). Measuring academic proficiency under the No Child Left Behind Act: Implications for educational equity. *Educational Researcher*.

Retrieved from

http://www.aera.net/uploadedfiles/publications/journals/educational_researcher/3408/01erv34n8_kim.pdf

King, M. B., & Bouchard, K. (2011). The capacity to build organizational capacity in schools.

Journal of Educational Administration, 49(5), 653-669.

Kovaleski, J.F., Gickling, E. E., & Marrow, H. (1999). High versus low implementation of

instructional support teams: A case for maintaining program fidelity. *Remedial*

Kulm, G. (2007). The achievement gap—Do we need a new strategy? *School Science*

and Mathematics, 107(2), 267-268.

Lee, V. E., Brooks-Gunn, J., Schnur, E., & Liaw, F. R. (1990). Are Head Start effects sustained?

A longitudinal follow-up comparison of disadvantaged children attending Head Start, No

preschool, and other preschool programs. *Child Development, 61*, 495–507. doi:

10.1111/j.1467-8624.1990.tb02795.x

Levin, B. (2007). Schools, poverty, and the achievement gap. *Phi Delta Kappan, 89*(1),

75-76.

Lewis, Ann L. (1982). An experimental evaluation of a direct instruction programme

(corrective reading) with remedial readers in a comprehensive school. *Educational*

Psychology, Vol. 2, No. 2, pp. 121–135.

Lindsey, J. (2010). *What the data really show: Direct instruction really works!* Retrieved

from: <http://www.jefflindsay.com/EducData.shtml>

Lodico, M.G., Spaulding, D.T., & Voegtler, K.H. (2010). *Methods in educational research*. New

York, NY: Harcourt Brace Jovanovich

Maccoby, E. E., & Zellner, M. (1970). *Experiments in primary education: Aspects of Project*

- Follow-Through*. New York, NY: Harcourt Brace Jovanovich.
- Mandinach, E. B., Honey, M., Light, D., Heinze, C., & Rivas, R. (2005, June). *Creating an evaluation framework for data-drive instructional decision making*. Paper presented at the meeting of the National Educational Computing Conference, Philadelphia, PA.
- Marzano, R. J., Pickering, D.J., & Pollock, J.E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: ASCD
- McLaughlin, M. W. (1995). *Improving education through standards-based reform*. Stanford, CA: National Academy of Education
- Meier, U. (2006). A note on the power of Fisher's least significant difference procedure
Department of Medical Data Services/Clinical Biostatistics, Boehringer Ingelheim
Pharma GmbH and Co. KG, Ingelheim am Rhein, Germany. meieru@ing.boehringer-
ingelheim.com.
- Mickelson, R .A. (2009). *School choice and segregation by race, class, and achievement*.
Retrieved from [http://nepc.colorado.edu/publication/school-choice-and-segregation-race-
class-and-achievement](http://nepc.colorado.edu/publication/school-choice-and-segregation-race-class-and-achievement)
- Monastersky, R. (2008). Researchers probe how poverty harms children's brains.
Chronicle of Higher Education, 54(25), A8.
- Myers-Miller, D. (2010). *Comparative evaluation of English language learner student achievements in twenty-one inner city urban schools*. (Doctoral dissertation). Retrieved from Dissertations & Theses: Full Text. (Publication No. AAT 3261845).
- National Association of Educational Statistics & U.S. Department of Education. (2011).
Findings in brief reading and mathematics 2011: National Association of Education

progress at grades 4 and 8. Retrieved from <http://nces.ed.gov/nationsreportcard/pdf/main2011/2012459.pdf>

National Center for Education Statistics. (2000). *Academic outcomes: Trends in the achievement gap in reading between black and white students*. Retrieved from nces.ed.gov/publis2000/coe2000/section2/indicator17.html

National Center for Fair and Open Testing. (2011). *No Child Left Behind annual report*. Retrieved from <http://fairtest.org/>

Neague, C. A. (2011). *Case study of effective instructional strategies that narrow the achievement gap*. (Doctoral dissertation). Retrieved from Dissertations & Theses: Full Text. (Publication No. AAT 345187).

Nehring, J. (2009). *The practice of school reform: Lessons from two centuries*. Albany, NY: SUNY Press.

Newman, F. (2007). *Answering the question that matters most: Has student achievement increased since No Child Left Behind?* Retrieved from: <http://www.cepd.org/document/docWindow.cfm?fuseaction=document.viewDocument&document id=200& document Formatted Id=3620>.

Noguera, P. A., & Akom, A. (2000, June 5). The causes of the racial achievement gap all derived from unequal treatment: Disparities demystified. *The Nation*. Retrieved from <http://www.highbeam.com/doc/1G1-62266830.html>.

O'Connor, M. (2001). Psychology responds to poverty. *Monitor Staff*, 32(9). Retrieved from: <http://www.apa.org/monitor/oct01/psychresponds.aspx>

Okoye, N. S. (2009). The effect of gender, socio-economic status, and school location on students' performance in Nigerian integrated science. *Education*, 129(2), 561-568.

- Orr, A. J. (2003). Black-white differences in achievement: The importance of wealth. *Sociology of Education*, 76, 281-304.
- Oseguera, L., Conchas, G. Q., & Mosqueda, E. (2011). Beyond family and ethnic culture: Understanding the preconditions for the potential realization of social capital. *Youth & Society*, 43(4), 1136-1166.
- Paris, K. (1994). *A leadership model for planning and implementing change for school-to-work Transition*. Madison, WI: University of Wisconsin-Madison, Center on Education and Work.
- Patten, M. D. (2011). Relationships between self-esteem, anxiety, and achievement in young learning disabled students. *Journal of Learning Disabilities*. 39(4), 255-331.
- Retrieved from <http://ldx.sagepub.com/content/16/1/43.abstract>
- Pearlstein, M. (2009). *Dwelling on institutional racism will not close the achievement gap*. Minneapolis: University of Minnesota.
-
- Pine, G. J., & Hilliard III, A. G. (1990). Rx for racism: Imperatives for America's schools. *Phi Delta Kappan*, 71, (6), 593-600
- Pogrow, S. (2002). Avoiding the stares when intellectually challenging disadvantaged students: Lessons from the HOTS programs. *Teachers.net Gazette*. Retrieved from <http://teachers.net/gazette/JAN02/pogrow.htm>
- Popham, J. W. (1975). *Educational evaluation*. Englewood Cliffs, NJ: Prentice-Hall.
- Porter, A. C. (2007). Lessons learned from standards-based reform: Implications for the reauthorization of NCLB. In A. Gamoran (Ed.), *Standards-based reform and the poverty gap: Lessons for "No Child Left Behind"* (p. 2). New York, NY: Brookings Institute.

- Porter, A. C. (2005). Prospects for school reform and closing the achievement gap. In C. A. Dwyer (Ed.), *Measurement and research in the accountability era* (p. 4). Mahwah, NJ: Lawrence Erlbaum.
- Present, W. (2010). Education reform in the United States and the impact of the No Child Left Behind Act. (Master's thesis). Retrieved from Dissertations & Theses: Full Text. (Publication No. AAT 1482329).
- Ractor, R., Johnson, K. & Fagan, P. (2001). *Understanding differences in black and white child poverty rates*. Retrieved from <http://www.heritage.org/>
- Rand Corporation. (2007). *Accountability under No Child Left Behind*. Retrieved from http://www.rand.org/pubs/research_briefs/RB9303/
- Ravitch, D. (2010). Welcome back to school “reform.” *Education Week Blog*. Retrieved from http://blogs.edweek.org/edweek/Bridging-Differences/2010/09/welcome_back_to_school_reform.html
- Ravitch, D. (2009, September 6). Time to kill No Child Left Behind. *Education Week*. Retrieved from: http://www.edweek.org/login.html?source=http://www.edweek.org/ew/articles/2009/06/04/33ravitch_ep.h28.html&destination=http://www.edweek.org/ew/articles/2009/06/04/33ravitch_ep.h28.html&levelId=2100
- Rawl, Ruth K. and O'Tuel, F. (1982). A comparison of three pre-reading approaches for kindergarten students. *Reading Improvement*, 19(3), 205–211.
- Reardon, S. F. (2011). *The widening academic achievement gap between the rich and the*

- poor: New evidence and possible explanations*. Stanford, CA: Stanford University.
- Retrieved from http://www.huffingtonpost.com/2011/11/21/income-achievement-gap-al_n_1105783.html
- Rist, R. C. (1973). *The urban school: A factory for failure. A study of education in American society*. Cambridge: MIT Press.
- Ritchey, F. J. (2000). *The statistical imagination: Elementary statistics for the social sciences*. New York, NY: McGraw-Hill.
- Robelen, E. W. (2008). Gates urges leaders to support education in troubled economy. *Education Week*, 28(15), 7.
- Rosenshine, B. (1996). Advances on research on instruction. *The Journal of Educational Research*, 88, 262-268.
- Rosenshine, B. (1976). *The seventy-fifth yearbook of the national society for the study of education* (Part 1). Chicago: University of Chicago Press.
- Ross, S. M., Nunnery, J. A., Goldfeder, E. McDonald, A., Rachor, R. (2004). Using school reform models to improve reading achievement: A longitudinal study of Direct Instruction and Success for All in an urban district. *Journal of Education for Students Placed at Risk*, 9(4), 357-388.
- Ross, S. M., Smith, L.J., Casey, J., Johnson, F., & Bond, C. (1994). *Using "Success for All" to restructure elementary schools: A tale of four cities*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
- Rothstein, R. (2008). Whose problem is poverty? *Educational Leadership*, 65(7), 8-13.
- Rothstein, R. (2011). *Richard Rothstein on the many causes of the achievement gap*. Harvard Education Letter. Retrieved from <http://www.hepg.org/page/112>.

- Rowley, R. L., & Wright, D. W. (2011). No "white" Child Left Behind: The academic achievement gap between black and white students. *Journal of Negro Education, 80*(2), 93-107.
- Rush, S. J. (2005). *Gender differences in reading and math scores on the Pennsylvania System of School Assessment grade five and grade eight tests*. (Doctoral dissertation).
- Ryder, R. J., Sekulski, J., & Silberg, A. (2003). *Results of direct Instruction reading program evaluation first through second grade 2000-2002*. Retrieved from [http://www.soe.uwm.edu/pages/welcome/Departments/Education Outreach/Programs__Courses/reading symposium/reading](http://www.soe.uwm.edu/pages/welcome/Departments/Education%20Outreach/Programs__Courses/reading%20symposium/reading)
- Ryder, R. J., Burton, J. L., & Silberg, A. (2006). Longitudinal study of Direct Instruction effects from first through third grade. *Journal of Educational Research, 99*(45), 179-192.
- Sadovnik, R. R., Cookson, Jr., P.W., & Semel, S.F. (2001). *Exploring education: An introduction to the foundations of education*. Needham Heights, MA: Allyn and Bacon.
- Sadowski, M. (2001). Closing the gap one school at a time: Teachers and administrators are becoming researchers as they work to narrow the black/white achievement gap in schools. *Harvard Educational Letter, 17*(3).
- Samuels, C. (2009). Response to intervention in math seen as challenging. *Education Week, 28*(35), 7. Retrieved from <http://www.linkedin.com/pub/3/751/65a>
- Sanders, M. G. (1997). Overcoming obstacles: Academic achievement as a response to racism and discrimination. *Journal of Negro Education, 66*(1), 83-93.
- Santrock, J.W. (2009). *Educational psychology* (4th ed.). New York, NY. McGraw Hill.
- Seaton, E. K. (2010). The influence of cognitive development and perceived racial

- discrimination on the psychological well-being of African-American youth. *Journal of Youth and Adolescence*, 39(2), 694-703.
- Slavin, R. L., & Madden, N.A. (1995). *Effects of Success for All on the achievement of English language learners*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Slavin, R. E., Madden, N.A., et al. (2005). *A randomized evaluation of success for all: Second-year outcomes*. Baltimore, MD: Johns Hopkins University, Center for Data-Driven Reform in Education.
- Smith, H. (2005). Making schools work. *PBS*. Retrieved from http://www.pbs.org/making_schools_work/index.html
- Speight, S. L. (2007). Internalized racism: One more piece of the puzzle. *Counseling Psychologist*, 35(1), 126-134.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape the intellectual identities. *American Psychologist*, 52, 613-629.
- Stern, S. (2009). E.D. Hirsch's curriculum for democracy: A content-rich pedagogy make better and smarter kids. *City Journal*, 19(4). Retrieved from http://www.city-journal.org/2009/19_4_hirsch.html
- Stockard, J. (2011). *Direct Instruction and first grade reading achievement: The role of technical support and time of implementation*. Retrieved from <http://www.nifdi.org/15/>
- Strayhorn, T. L. (2010). When race and gender collide: Social and cultural capital's influence on the academic achievement of African-American and Latino males. *Review of Higher Education*, 33(3), 307-332.
- Suizzo, M. A., Robinson, C., & Pahlke, E. (2008). African-American mothers'

- socialization beliefs and goals with young children: Themes of history, education, and collective independence. *Journal of Family Issues*, 29(3), 287-316.
- Sun, Y., & Li, Y. (2011). Effects of family structure type and stability on children's academic performance trajectories. *Journal of Marriage and Family*, 73, 541-556.
- Swim, J. K., Hyers, L. L., et al. (2003). African American college students' experiences with everyday racism: Characteristics of and responses to these incidents. *Journal of Black Psychology*, 29(2), 38-67.
- Tan, E., & Barton, A. C. (2010). Transforming science learning and student participation in sixth grade science: A case study of a low-income, urban, racial minority classroom. *Equity & Excellence in Education*.43 (1), 38-55.
- Tarver, S. G. (2004, February 25). Letter to the editor: Direct instruction—Criticism of a Wisconsin study. *Education Week*, 23(24), 38.
- Taylor, J. A. (2005). Poverty and student achievement. *Multicultural Education*, 12(4), 56-57.
- Thattai, D. (2008). A history of public education in the United States. Retrieved from <http://www.servinfree.net>
- Thompson, B. (2007). Effect sizes, confidence intervals, and confidence intervals for effect sizes. *Psychology in the Schools*, 44, 423-432.
- Thompson, B., Diamond, K.E., et al. (2005). **Correlational** research for evidence-based practice. Retrieved from: <http://www.cec.sped.org/Content/NavigationMenu/ProfessionalDevelopment/ProfessionalStandards/Copy%20of%20Final%20Manuscript-Correlational%20Research%204-6-04.pdf>
- Thompson, M., & Thompson, J. (2000). *Leadership, achievement, and accountability:*

- Benchmarking to exemplary practice*. Boone, NC: Learning Concepts.
- Toppo, G. (2009, March 5) What to learn: 'Core knowledge' or '21st-century skills'? *USA TODAY, News Education*. Retrieved from:
http://www.usatoday.com/news/education/2009-03-04-core-knowledge_N.htm
- Trochim, M. K., & Donnelly, J. P. (2008). *Research methods knowledge base*. Cleveland, OH: Cengage Learning Center.
- U.S. Census Bureau. (2008). *Educational attainment by selected characteristics: 2006*. N.Y. Retrieved from http://www.census.gov/compendia/statab/cats/education/educational_attainment.html
- U.S. Census Bureau. (2009). *Poverty thresholds for 2009 by size of family and number of related children under 18 years*. Retrieved from <http://www.census.gov/hhes/www/poverty/data/threshold/thresh09.html>
- U.S. Census Bureau. (2011). *Poverty, social, economic, and housing statistics division*. Retrieved from: <http://www.census.gov/hhes/www/poverty/data/threshld/index.html>
- U.S. Department of Agriculture. (2011). *Determining and verifying eligibility: Child nutrition programs food and nutrition services*. Retrieved from <http://www.fns.usda.gov/cnd/Guidance/EliMan.pdf>
- U.S. Department of Education. (n.d.). *No Child Left Behind*. Retrieved from <http://www.ed.gov/nclb/landing.jhtml?src=pb>
- U.S. Department of Education. (1983). *A nation at risk*. Retrieved from <http://www2.ed.gov/pubs/NatAtRisk/risk.html>. U.S. Department of Education. (2002). *Scientifically-based research and the comprehensive school reform (CSR) program*.

- Washington, DC: Author. Retrieved from
<http://www.ed.gov/offices/OESE/compreform/appendc.pdf>
- U.S. Department of Education. (2007a). *IDEA regulations: Alignment with the No Child Left Behind*. Washington, DC: Author.
- U.S. Department of Education. (2007b). *Title I — Improving the academic achievement of the disadvantaged*. Washington, DC: Author.
- U.S. Department of Education. (2008). *President's budget strengthens nation's commitment to No Child Left Behind*. Washington, DC: Author.
- U.S. Department of Education. (2010). *The nation's report card*. Washington, DC: Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP). Retrieved from
<http://nces.ed.gov/nationsreportcard/pdf/main2011/2012459.pdf>
- United States Educational Policy. (2008). *The basics of educational policy; The pressure for reform in American education, defining policy*. Retrieved from:
<http://education.stateuniversity.com/pages/1937/Educational-Policy-Unted-States.html>
- Van Dorn, B. (n.d.). The core knowledge curriculum. *Family Education*. Retrieved from
<http://school.familyeducation.com/curriculum-planning/alternative-education/38291.html#ixzz1eStwtyY9>
- Vaughn-Purdie, V. (2004). *Combating the negative effects of stereotypes: Improving minority performance with a values-affirmation intervention*. Retrieved from
<http://www.columbia.edu/cu/psychology/fac-bios/PurdieVaughnsV/faculty.html>
- Venezky, R.L. (1994). *An evaluation of Success for All: Final report to the France and Merrick Foundations*. Newark, DE: University of Delaware, Department of Educational Studies.

- Viadero, D. (2010). *Federal school reform program failed to pan out*. Retrieved from:
http://blogs.edweek.org/edweek/insideschoolresearch/2010/04/study_says_federal_csr_program.html
- Volonino, V., & Zigmond, N. (2007). Promoting research-based practices through inclusion?
Theory Into Practice, 46, 291-300.
- Warner, R. M. (2008). *Applied statistics: From bivariate through multivariate techniques*. Thousand Oaks, CA: Sage.
- Watkins, T. (2008). *A comparative analysis of the effectiveness of Direct Instruction reading on African-American, Caucasian, and Hispanic students*. (Unpublished doctoral dissertation). Delta State University, West Cleveland, MS.
- Webber, C. (2001). Racism and the achievement gap. *Education Week, 20*(43), 49-72.
- Wenglinsky, H. (2004). Closing the racial achievement gap: The role of reforming instructional practices. *Education Policy Analysis Archives, 12*(64). Retrieved from <http://epaa.asu.edu/epaa/v12n64/>
- West, C. (2001). *Race Matters*. Vintage Books. NY.
- White, W. A. T. (1988) Meta-analysis of the effects of direct instruction in special education. *Education and Treatment of Children, 11*, 364-374.
- Wiersma, W., & Jurs, S. G. (2005). *Research methods in education: An introduction*. Boston, MA: Bacon.
- Williams, A. (2011). A call for change: Narrowing the achievement gap between white and minority students. *Clearing House: A Journal of Educational Strategies, Issues, and Ideas, 84*(2), 65-71.
- Williams, A. (2006). *No Child Left Behind Policy: Is reform still necessary?* Washington, DC:

U.S. Army War College.

Woestehoff, J. (2011). Arne Duncan “reforms” increased black-white achievement gap.

Examiner.com. Retrieved from <http://www.examiner.com/k-12-in-chicago/arne-duncan-reforms-increased-black-white-achievement-gap>

Woodson, C. G. (1990). *The mis-education of the Negro*. Trenton, NJ: Africa World Press.

Woolfolk, A. (2010). *Educational psychology* (11th ed.). Upper Saddle River, NJ: Pearson.

Yeung, W. J., & Conley, D. (2008). Black-White achievement gap and family wealth. *Child Development, 79*, 303-324.

Yudof, S., & Leonard, E. (2008). *U.S. Secretary of Education Margaret Spellings marks 25th anniversary of A Nation at Risk report by broadening the national dialogue on the state of American education*. Washington, DC: U.S. Department of Education Office of Communications.

APPENDIX A: IRB APPROVAL



The Graduate School at Liberty University

February 15, 2013

Eric Fuller

IRB Exemption 1539.021513: The Impact of Selected Research-based Instructional Reform Programs, Success for All (SFA), Direct Instruction (DI) and Core Knowledge on the CRCT Reading Scores of African-American Students and Students from Disadvantaged Subgroups in Grades Three, Five, and Eight in a Large Urban School System

Dear Eric,

The Liberty University Institutional Review Board has reviewed your application in accordance with the Office for Human Research Protections (OHRP) and Food and Drug Administration (FDA) regulations and finds your study to be exempt from further IRB review. This means you may begin your research with the data safeguarding methods mentioned in your approved application, and that no further IRB oversight is required.

Your study falls under exemption category 46.101 (b)(4), which identifies specific situations in which human participants research is exempt from the policy set forth in 45 CFR 46:

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

Please note that this exemption only applies to your current research application, and that any changes to your protocol must be reported to the Liberty IRB for verification of continued exemption status. You may report these changes by submitting a change in protocol form or a new application to the IRB and referencing the above IRB Exemption number.

If you have any questions about this exemption, or need assistance in determining whether possible changes to your protocol would change your exemption status, please email us at irb@liberty.edu.

Sincerely,

Fernando Garzon, Psy.D. Professor, IRB Chair **Counseling (434) 592-4054** Liberty University

APPENDIX B: TABLES ON NUMBER OF STUDENTS TESTED

Table B1

Number of Students Tested: African-American Students and Disadvantaged Students by School and Percent Meeting or Exceeding Standards

School	School ID	School Type	Number Tested	Reading % met or exceeded standard
--------	-----------	-------------	---------------	------------------------------------

Note. R=schools with no whole-school reform. NR=schools with whole-school reform

School	1	R	54	92.6
School	2	NR	64	68.8
School	3	NR	67	74.6
School	4	R	74	77.0
School	5	N	85	88.2
School	6	NR	54	100.0
School	7	NR	36	75.0
School	8	NR	81	85.2
School	9	NR	41	92.7
School	10	R	67	80.6
School	11	R	57	98.2
School	12	NR	80	100.0
School	13	R	42	61.9
School	14	NR	64	82.8
School	15	NR	10	40.0
School	16	NR	60	76.7
School	17	NR	58	81.0
School	18	R	118	83.1
School	19	NR	30	90.0
School	20	NR	36	88.9
School	21	NR	56	100.0
School	22	R	69	72.5
School	23	NR	78	94.9
School	24	R	49	83.7

School	School ID	School Type	Number Tested	Reading % met or exceeded standard
--------	-----------	-------------	---------------	------------------------------------

Note. R=Schools with whole-school reforms. NR=Schools with no whole-school reforms

School	25	R	57	86.0
School	26	NR	67	77.6
School	27	R	82	74.4
School	28	NR	80	91.3
School	29	NR	47	83.0
School	30		48	91.7
School	31	NR		
	31		85	83.5
School	32	NR	144	99.3
School	33	R	96	99.0
School	34	R	60	70.0
School	35	NR	123	99.2
School	36	R	83	69.9
School	37	NR	37	89.2
School	38	R	30	86.7
School	39	NR	84	70.2
School	40	NR	126	97.6
School	41	NR	94	96.8
School	42	NR	69	71.0
School	43	R	43	100.0
School	44	NR	41	82.9
School	45	NR	41	70.7
School	46	R	41	75.6
School	47	NR	86	82.6
School	48	NR	49	89.8
School	49	R	98	82.7
School	50	NR	58	86.2

Table B3

Number of Students Tested: African-American Students and Disadvantaged Students by School and Percent Meeting or Exceeding Standards

School	School ID	School Type	Number Tested	Reading % met or exceeded standard
School	52	NR	43	90.7
School	53	R	42	83.3
School	54	R	87	86.2
School	55	NR	86	67.4
School	56	NR	26	100.0
School	57	R	39	97.4
School	58	NR	42	76.2
School	59	R	52	80.8
School	60	NR	45	62.2
School	61	NR	75	66.7
School	62	NR	41	97.6
School	63	NR	56	67.9
School	64	R	24	91.7
School	65	NR	62	72.6
School	66	R	183	94.0
School	67	NR	165	90.3
School	68	NR	163	90.8
School	69	R	268	91.8
School	70	NR	109	93.6
School	71	NR	91	94.5
School	72	R	48	97.9
School	73	NR	12	66.7
School	74	NR	256	93.0

Table B4

Number of Students Tested: African-American Students and Disadvantaged Students by School and Percent Meeting or Exceeding Standards

School	School Identifier	School type	Number tested	Reading % met or exceeded standard
School	75	NR	15	87.3
School	76	R	185	90.8
School	77	NR	90	71.1
School	78	R	49	98.0
School	79	NR	188	84.6
School	80	NR	72	98.6
School	81	R	22	100.0
School	82	NR	115	87.0
School	83	NR	244	98.4
School	84	R	329	97.9
School	85	NR	107	93.5
School	86	NR	94	92.6
School	87	NR	125	88.0