The Relationship between Participation in Extended-Day Programs and Mathematics Achievement of Title I-Eligible Students

Crystal M. Lovell
Seton Hall University

Follow this and additional works at: https://scholarship.shu.edu/dissertations
Part of the Educational Administration and Supervision Commons, and the Science and Mathematics Education Commons

Recommended Citation
https://scholarship.shu.edu/dissertations/1750
THE RELATIONSHIP BETWEEN PARTICIPATION IN EXTENDED-DAY PROGRAMS AND MATHEMATICS ACHIEVEMENT OF TITLE I-ELIGIBLE STUDENTS

BY

CRYSTAL M. LOVELL

Dissertation Committee

Charles M. Ackilias, Ed.D. Mentor
Anthony J. Colella, Ph.D.
Henry L. Allen, Ph.D.
Marylu Simon, Ed.D.

Submitted in Partial Fulfillment of the Requirements for the Degree Doctor of Education Seton Hall University

2006
ABSTRACT

The Relationship between Participation in Extended-day Programs and Mathematics Achievement of Title I-Eligible Students

Background: In 1965, the Federal government declared war on poverty and the educational inequities resulting from the stratification of the socioeconomic classes. Title I provided educational resources to raise achievement and close the socioeconomic status achievement gap. Research on the effectiveness Title I implementation remains inconclusive.

Purpose: To assess the relationship of participating in a Title I-funded extended-day mathematics program and improving academic achievement and closing the minority and SES achievement gaps.

Setting: Suburban district in Central New Jersey.

Population: 284 students enrolled in Grades 2 through 5.

Intervention: 87 Title I-eligible students received 60-75 minutes of mathematics instruction after school weekly. Non-participate comparison groups included 75 Title I-eligible and 122 ineligible students.

Research Design: Nonexperimental, cross-sectional, explanatory research design.

Data: t-test analyses of DBA mathematics pre- and posttests, TerraNova percentile rankings, and NJ ASK 3 and 4 assessments. Information regarding instructional strategies used, nature of staff training, content taught, time-on-task, curriculum materials used and class size were collected from teacher anecdotal notes.
**Findings:** Although participants scored significantly lower on the DBA pre- and posttests than did the comparison groups, there was no significant difference in the pre-posttest gain scores between the groups. Participants were progressing at the same rate as all other students. Participants scored significantly lower on 6 of the 7 NJ ASK indicators than did the comparison groups, however the NJ ASK mathematics cluster reliability estimates are relatively low and the standard errors of measurement are relatively high. There was no significant difference between participant performance and that of the comparison groups on 3 of 7 TerraNova indicators. Participants scored significantly higher on 1 of the 7 TerraNova indicators than did the comparison groups.

**Conclusions:** Findings are congruent with current research which shows that short-term participation in Title I extended-day programs yield small positive gains at improving achievement and few, if any, positive gains at closing the achievement gaps. However, similar to class size research, the findings show that long-term, consecutive-year participation in Title I extended-day programs that incorporate and maximize the opportunity-to-learn (OTL) variables yield positive gains at improving academic achievement and closing the achievement gaps.
DEDICATION

My husband’s favorite Bible verse is “I can do all things through Christ who strengthens me” Philippians 4:13. It seems appropriate that this accomplishment be a testament to Him. I am most grateful to God for giving me the gifts and talents to pursue this dream, providing me with guidance throughout my journey, and giving me stamina and perseverance to complete this accelerated doctoral program. God is Good – All the Time!

Second to God is my family – to whom this is dedicated. I am blessed to have a terrific, loving family: my husband, Gerard, my children, Robyn, Trey and Miracle, my mom, Jean, and my sister Jeanell.

Gerard – Thanks for supporting, inspiring and believing in me even when I didn’t. Although you think I don’t always listen to your words of wisdom, I really do – I just don’t tell you.

Robyn – Thanks for being “me” for the last two years. You were a mother to your sisters while I attended school – cooking, cleaning, entertaining and helping them with their homework. I couldn’t ask for a more special gift than to have you as my daughter.

Trey and Miracle – Thanks for the back rubs, foot massages, and hundreds of kisses that kept me going. Mommy loves you.

Mom – Words can not express all that you have done for me to make this possible. I hope I can be as good of a mom to my kids as you are to me.

Jeanell – Thanks for being a supportive sister, a substitute mom, and a good friend. I wouldn’t be in the doctoral program if you didn’t help me study for the MAT.
I wish to thank my mentor, Dr. Charles M. Achilles for all of his time, guidance and support during the last two years. You made me work hard and pushed me beyond what I thought I was capable of doing; however, in the end it was well worth it and I am a better educator because of it.

Dr. Anthony J. Colella and Dr. Henry L. Allen – Thank you for your patience and assistance. Despite your busy schedules you always had time to read my draft copies and provide feedback in a timely manner.

Dr. Marylu Simon – Thank you for inspiring me and supporting me in achieving this advanced degree. You have helped me accomplish so much more than just this degree.

Cohort VIII - I could not ask for a better support system than this group. You are truly a magnificent group of people and I am a better person for having spent two intense years working with you.

Lauren Spicer, Carolyn Ash and Paula Childs – Thanks for being my prayer warriors and spiritual sisters.

Brian Purvack and Freddie Evans – Thanks for keeping me going when the times got rough.

Kyle and Angélique McGlotten – Thanks for keeping me in your prayers. Friends like you are priceless.
TABLE OF CONTENTS

LIST OF TABLES ............................................................................................................. viii
LIST OF FIGURES ........................................................................................................... ix

I. INTRODUCTION ........................................................................................................... 1
   Introduction ................................................................................................................... 1
   History of Title I ......................................................................................................... 4
   Title I programs .......................................................................................................... 12
   The Research Problem .............................................................................................. 14
   Purpose of the Study .................................................................................................. 18
   Delimitations and Limitations .................................................................................. 19
   Research Questions .................................................................................................... 22
   Definition of Terms ..................................................................................................... 24
   Summary of Chapter One and Organization of Remainder of Study ..................... 27

II. LITERATURE AND RESEARCH REVIEW .................................................................... 29
   Introduction .................................................................................................................. 29
   Historical Reviews ...................................................................................................... 31
   Efforts to Improve Academic Achievement ............................................................. 32
   Efforts to Close Achievement Gaps ........................................................................... 37
   Other Related Factors ................................................................................................. 38
   Theoretical framework ............................................................................................... 38
   Time ............................................................................................................................. 39
   Instructional practices ............................................................................................... 44
   Class Size ..................................................................................................................... 46
   Staffing and staff training .......................................................................................... 49
   Curriculum ................................................................................................................... 51
   Program evaluation ..................................................................................................... 53
   Opportunity to learn .................................................................................................... 54
   Summary ....................................................................................................................... 56

III. METHODOLOGY ......................................................................................................... 58
   Introduction .................................................................................................................. 58
   Research Design and Variables .................................................................................. 59
   Participants .................................................................................................................... 64
   Instrumentation .......................................................................................................... 68
   The Program ................................................................................................................ 74
   Data Collection and Analysis ..................................................................................... 77
   Hypotheses ................................................................................................................... 98
   Summary ....................................................................................................................... 101
IV. DATA AND ANALYSES

Introduction .................................................................................................................. 103
Research Question One ............................................................................................... 106
Summary of Research Question One .......................................................................... 108
Research Question Two ............................................................................................... 108
  Comparison of Project Excel participants and all non-participants .................... 110
  Comparison of Project Excel participants to district, state, and national norms .... 113
  Comparison of Project Excel participants and Title l-eligible non-participants ..... 115
Summary of Research Question Two ......................................................................... 120
Research Question Three ........................................................................................... 121
  Minority achievement gap ...................................................................................... 121
  SES achievement gap ............................................................................................. 126
Summary of Research Question Three ..................................................................... 127
Additional Findings .................................................................................................... 133
  Class size and attendance ....................................................................................... 133
  Time, methodology and curriculum ...................................................................... 134
  Staffing and staff training ....................................................................................... 135
Summary ....................................................................................................................... 137

V. SUMMARY OF FINDINGS, CONCLUSIONS, DISCUSSION, AND RECOMMENDATIONS

Introduction .................................................................................................................. 139
Summary of Findings .................................................................................................. 141
  Research question one ............................................................................................ 141
  Research question two ............................................................................................ 142
  Research question three ........................................................................................ 144
Conclusions and Discussion ....................................................................................... 146
Recommendations for Practice, Policy and Future Research ................................. 158
  Longitudinal cohort analyses ............................................................................... 158
  Program evaluation ............................................................................................... 160
  Early intervention ................................................................................................. 161
  Mandatory participation ....................................................................................... 162
  Small class size ..................................................................................................... 163
  Extended-year programs ....................................................................................... 165
  Future research ..................................................................................................... 166
Summary ....................................................................................................................... 170

References .................................................................................................................... 172

Appendix A: Everyday Mathematics Learning Goals .................................................. 187
<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Study Sample by Grade and Group Assignment</td>
<td>65</td>
</tr>
<tr>
<td>2</td>
<td>Demographics of Student Population</td>
<td>67</td>
</tr>
<tr>
<td>3</td>
<td>Final Study Sample - Special Education and G&amp;T Students Excluded</td>
<td>69</td>
</tr>
<tr>
<td>4</td>
<td>Reliability Estimates and Standard Errors of Measurement (SEM) for Mathematics Clusters – 2004 NJ ASK Grades 3 and 4</td>
<td>73</td>
</tr>
<tr>
<td>5</td>
<td>Reliability Estimates and Standard Errors of Measurement (SEM) for TerraNova Spring Form A Multiple Assessments</td>
<td>75</td>
</tr>
<tr>
<td>6</td>
<td>Summary of Subsidiary Questions and Statistical Analyses Associated with Research Question Two</td>
<td>88</td>
</tr>
<tr>
<td>7</td>
<td>Summary of Subsidiary Questions and Statistical Analyses Associated with Research Question Three</td>
<td>94</td>
</tr>
<tr>
<td>8</td>
<td>DPA Pre-and posttest Gains for Second and Third-Grade Students</td>
<td>132</td>
</tr>
<tr>
<td>9</td>
<td>Most Common instructional Methodologies used by Project Excel Instructors</td>
<td>134</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1 Theoretical framework for the present study ........................................40
2 Graphical representation of the first research question .................... 79
3 Graphical representation of the second research question ................. 82
4 Graphical representation of the third research question .................... 91
5 Mean DBA mathematics pretest scores for Groups A, B, B and C, and C ................................................................... 107
6 Mean DBA mathematics pretest scores categorized by race and SES .... 109
7 Mean DBA mathematics posttest scores for participants and non-participants ........................................................................ 111
8 Mean TerraNova mathematics percentile rankings for participants and non-participants ........................................................................ 112
9 Mean NJ ASK mathematics scores for participants and non-participants ..................................................................................... 114
10 Mean DBA mathematics posttest scores for participants and eligible non-participants ........................................................................ 116
11 Mean mathematics TerraNova percentile rankings for participants and eligible non-participants ........................................................................ 118
12 Mean mathematics NJ ASK scores for participants and eligible non-participants ..................................................................................... 119
13 Mean DBA mathematics posttest scores for non-Asian minority and White non-participants ........................................................................ 122
14 Mean TerraNova mathematics percentile rankings for non-Asian minority and White non-participants ........................................................................ 124
15 Mean NJ ASK mathematics scores for non-Asian minority and White non-participants ........................................................................ 125
16 Mean DBA pre- and posttest gains for Groups A, B, B and C, and C ..................................................................................... 130
17 Mean DBA pre- and posttest gains categorized by race and SES .... 131
18 Analysis of average Project Execs instructional time spent teaching various mathematical concepts .................................................... 136
Chapter I

Introduction

Introduction

The National Center for Children in Poverty (2004) reported that in 2002 over 26 million children in the United States were growing up in low-income families. That figure represented about 37% of the children in the country. While the largest population of children was White, 58% of all African American children and 62% of all Latino children were growing up in low-income families. How do the conditions associated with poverty relate to student achievement in school?

Barton (2004) reported 14 life experiences and conditions associated with low-income status that were also associated with student achievement. Conditions such as low birth weight, lead poisoning, hunger and malnutrition were some of the factors that delayed motor, cognitive and social development. In a similar study Rothstein (2004b, 2005) identified conditions such as improper vision, hearing and oral health care that were often associated with low-income families were also associated with lower student achievement.

In addition to the health related factors, Rothstein (2004a, 2004b, 2005) and Barton (2004) reported that the degree of parent-child interaction is associated student achievement. Language acquisition skills and literacy development are related to how often young children are read to and the types of interactions the parents have with the children while reading. Rothstein (2004b) reported that while reading to their children
"working-class parents are more likely to tell children to pay attention without interruptions or to sound out words or name letters...[while] [p]arents who are more literate are more likely to ask questions that are creative, interpretive or connective" (p. 21). Children raised in the latter environment tend to read more for pleasure and enjoyment. Hart and Risley as cited in Rothstein (2004a) stated "[o]n average, professional parents spoke more than 2,000 words per hour to their children, working-class parents spoke about 1,300 and welfare mothers spoke about 600" (p. 41). Hart and Risley as summarized in Rothstein (2004a) also noted that professional parents were six times more likely to encourage their children rather than reprimand them, while the welfare parents were two times more likely to reprimand their children rather than encourage them. By the time they reached school-age, children of professional parents had a vocabulary that consisted of more than three times as many words than did the vocabulary of children of welfare parents. Rothstein (2004b) suggested that not only did the children of professional parents have a larger vocabulary than their welfare counterparts, these children had a larger vocabulary than the welfare parents.

The Child Trends Data Bank as cited in Barton (2004) showed that parent availability and participation in school-related activities can be associated with fewer student behavioral problems and better academic performance. "Most parents attend scheduled meetings with teachers, but parents of Black and Hispanic students and low-income parents are much less likely than parents of White students are to attend a school event, do volunteer work, or serve on a committee" (Barton, 2004, p. 11).

Rothstein (2004b) suggested that in addition to the decrease in participation in school-sponsored activities, low-income parents were more likely to change schools often.
due to lack of affordable housing. Similarly, Barton (2004) noted that increased student mobility was associated with below grade performance in reading and math.

Watching too many hours of television is also associated with lower achievement, and Barton (2004) reported that 42% of Black fourth-grade students watched 6 hours or more of television each day. Christakis as cited in Barton (2004) "established that each hour of television that a child watches on a daily basis between the ages of 1 and 3 years old increases by 10 percent the risk that the child will have attention problems" (p. 10). Television is often used as a babysitter in low-income homes.

Most of the aforementioned conditions occurred outside the school doors and existed well before the child began formal schooling; however, the researcher sought to evaluate school-related factors that are associated with student achievement. The rigor of the curriculum; experience, preparation and attendance of the teacher; class sizes; level of technology-assisted instruction; and the safety of the learning environment were a few additional factors that Barton (2004) noted that were associated with student achievement, and these factors can be controlled by school district leaders. So what role does the American public school system have in combating the ill-effects of poverty? Clayton (1991) stated:

For well over a century, a commitment to educational opportunity through public schooling for all children has been a basic premise of our American democracy...[and] that premise was strengthened by the articulation of the concept of compensatory education; Schools can and should compensate for the effects of poverty on children's lives by providing them with extra support needed to improve their life chances. (p. 345)
Since the inception of Title I, school districts have reportedly spent millions of dollars on supplemental and compensatory educational programs with the hope of improving academic achievement and closing the achievement gaps. Although such an investment is an important endeavor, Rothstein (2005) stated that “[i]t is foolish for anyone to think that educational reform is the solution to closing the achievement gap...unless we eliminate the social and economic inequities that exist in this society, the achievement gap will persist” (Speech presented at an Educational Testing Service conference). Yet Rothstein was not against school reform; rather, he supported initiatives that would combine economic and social reform with educational reform. “School improvement does have an important role to play, but it cannot shoulder the entire burden, or even most of it, on its own” (Rothstein, 2004b, p. 149). However, history will show that lawmakers repeatedly looked to schools as the sole solution for closing the achievement gap. What originally began as financial support to assist impoverished children has changed into a political arena loaded with accountability systems such as No Child Left Behind (NCLB) and penalties for not achieving the federal government-defined Adequate Yearly Progress (AYP).

History of Title I. In 1965, Congress initiated Title I of the Elementary and Secondary Education Act to provide educational resources for impoverished children. Taylor and Teildicke (1999) stated that the goal of Title I “was to raise achievement levels for these youngsters in reading, writing, and mathematics to levels approximating those of more advantaged children” (p. 299). So what is the status of American education over 40 years after the inception of ESEA? The U.S. Department of Education (2004a) noted that:
A significant achievement gap [still] exists between disadvantaged students and their more affluent peers, despite billions in Federal spending since 1965... Since the Elementary and Secondary Education Act first passed Congress in 1965, the federal government has spent more than $321 billion dollars [in 2002] to help educate disadvantaged children. Yet nearly 40 years later, only 32 percent of fourth-graders can read skillfully at grade level. Sadly, most of the 68 percent who can't read well are minority children and those who live in poverty. (U.S. DOE, PowerPoint slide 7)

Initially Title I funds were available to schools in which 75% or more of their student population was defined as low-income. There were few limitations of how the money could be spent, and even fewer accountability systems that existed to keep track of the money. The major requirement was that funding provided specific services only to those impoverished students who qualified to receive them (Wong & Meyer, 1998). With such loose regulations and an influx of over a billion dollars of new money, some schools saw this as an opportunity to do more than simply educate the disadvantaged. Cook (2005) reported that “[s]candals erupted in the late 1960s and early 1970s over schools purchasing audiovisual equipment and other supplies with Title I dollars” (p.25). Congress responded by tightening the rules in the early 1970s, requiring districts to rank their schools according to need and provide matching funds.

District leaders saw the matching funds as a punitive measure and responded by creating many inferior programs that often isolated the children. These services consisted of pull-out or pull-aside programs, where children were either removed from the
classroom or moved to a private location within the classroom to receive separate instruction. Such programs received criticism as being disruptive, causing students to miss instruction with their regular teacher (Le Tendre as cited in Taylor & Teddie, 1999). Schools were criticized for the disparities in the curriculum and instruction Title I students received because these students often received different texts, tests, curriculum materials and were frequently taught by noncertified teachers and those lacking experience with effective instructional strategies (Cook, 2005). "Research began to document the disadvantages of these strategies, including the negative impact on students and the contribution of pullout programs to curriculum and instructional fragmentation" (Birman et al.; Knapp, Shields & Turnbull, as cited in Sunderman, 2001, p. 506).

Congress responded to school district concerns by amending Title I legislation in 1978 making it possible for educators to integrate Title I and regular curricular programs, thereby creating a schoolwide program approach (Taylor & Teddie, 1999). School district leaders were slow to implement schoolwide programs because the law continued to require districts to match federal grants with their own funds. Hence, in many districts the practice of pull-out programs continued despite the disadvantages identified by the research.

During the Reagan era of the 1980s, Title I received a same change to Chapter I, and many of the monitoring and paperwork requirements were reduced, but the prevalence of the matching funds stipulation continued to act as a barrier to eliminating the pull-out and isolation programs. By 1988, Congress instituted the Hawkins-Stafford Amendment to ESEA which limited some restrictions to allow more flexibility of where and how to use Title I funds (Sunderman, 2001). The Hawkins Amendment removed the
matching funds requirement allowing educators to depart from the reliance of pullout programs and investigate the implementation of schoolwide programs. The Amendment also allowed districts to use Title I funds for:

1. Hiring additional staff to reduce class size and strengthen the relationship between the schools and families (Wong, 2003);

2. Facilitating district activities to promote parental involvement (Wong, 2003);

3. Implementing or significantly strengthening teacher training in reading/language arts and mathematics instruction (Schenck & Beckstrom, as cited in Wong, 2003);

4. Significantly strengthening components related to curriculum and instruction, such as computer-assisted instruction, supplemental education; provisions for extended school days and programs such as Reading Recovery and Success for All (Schenck & Beckstrom as cited in Wong, 2003);

5. Adopting practices associated with effective schools, including needs assessment, staff development, changes in classroom instruction, and changes in school management (U.S. DOE as cited in Wong, 2003).

The ESEA underwent another change in 1994 when Congress introduced the Improving America's Schools Act (IASA). The IASA mandated that Title I children be held to the same academic standards reflected in testing as all other students (Olson, 2001; Sunderman, 2001). To qualify for Title I funds states were required to develop content and performance standards, to adopt or develop student assessments that were
aligned with the standards, and to use the assessments as criteria to determine the progress of schools toward meeting the goal of improving student performance (Sunderman, 2001; Wong & Meyer, 1998). The LASA also lowered the eligibility threshold for Title I from a minimum of 75% low-income enrollment to a minimum of 50% low-income enrollment (Le Tendre & Chabran, 1998; Sunderman, 2001; Wong, 2003). The combined lowering of the threshold and removal of the matching funds requirement resulted in a dramatic increase on the number of schoolwide programs. Wong and Meyer (1998) stated that:

As a result of these federal legislative changes, the number of schoolwide programs grew from 1,200 in 1991 (of approximately 12,000 eligible schools) to over 8,000 in 1996 (of approximately 16,800 eligible schools), representing an increase from about 10 to 50% of the eligible schools in each of those years. (p. 116)

Title I was at the forefront of partisan politics in the mid 1990s. Politicians began to scrutinize the requirements for funding and investigated possible outcomes-based results as a necessary requirement for continued funding. Wong (2003) proposed that from the political arena emerged three possible courses for the future of Title I. The first direction was a call for adoption of externally designed models that had proven effectiveness elsewhere. The research in this area was scant as there was a lack of hard evidence to support adopting a program without considering the impact of local circumstances on feasibility.

The second direction was to invoke sanctions that required districts and schools to annually raise academic standards and improve professional development. Politicians
felt a strong performance-based accountability system needed to be in place. Using vouchers as a means of reforming Title I was the third option. Parents would choose whether schools were meeting the needs of their children, and federal money would follow the child to the school of choice. Politicians hoped the competition for funds would pressure schools to improve.

In 1999 the U.S. Department of Education worked in cooperation with the Council of Chief State School Officers (CCSSO) to establish an annual report series that provided individual state profiles of multiple indicators of compliance with IASA (Blank & Manise, 2000). The CCSSO School Indicators Report served as the watchdog, putting the actions and accountability of the states in the forefront of the public eye. It listed how many states had established content standards that met review criteria, constructed state assessments linked to standards, established minimum levels of proficiency, and reported student achievement results disaggregated. The report was presented to a variety of audiences and served to educate the general public on the performance of schools and their progress in improving student mathematics achievement. This was the beginning of the disaggregated data era in which data about student demographics and academic achievement were reported on a state level. Later it became mandated that all districts and schools report disaggregated student achievement data.

In 2002 Congress adopted No Child Left Behind (NCLB) legislation which had a significant impact on getting states to comply with previous mandates of the ESEA (U.S. Department of Education, 2004a). NCLB encompassed all three options discussed by Wong (2003): (a) implementation of scientifically research-based programs of instruction, (b) school accountability through AYP mandates, and (c) school choice
through sanctions for failing to meet AYP for 3 or more consecutive years. Prior Title I legislation focused on improving academic performance of economically disadvantaged students, while current NCLB legislation supports accountability measures to ensure academic progress of subgroups of students. Economically disadvantaged students make up only one of 40 subgroups which include students who are Limited English Proficient, students with disabilities, migrant students, male students, female students, and various ethnic/racial groups of students. A school whose overall performance was satisfactory may actually fail to reach NCLB standards if one or more of its subgroups did not demonstrate adequate yearly progress (AYP). Schools were often plagued by double jeopardy as many of their students were counted multiple times as representative members of underperforming subgroups. Consider the impact of the performance of an economically-disadvantaged, special education student whose native language is Spanish. States were given a deadline of 12 years to ensure that 100% of all students were proficient. In a memo submitted to editorial writers on March 11, 2004, U.S. Secretary of Education, Ronald Paige wrote:

For the first time in history, every state has an approved accountability plan to ensure academic proficiency for every child. Achievement gaps are being identified and addressed. The success of schools is now being measured on the academic achievement of all students so that children who need help aren't hidden in averages. (p.1)

Clayton (1991) pointed out that "[k]ey concerns of the Chapter I [Title I] legislation are ensuring that the services provided to students are producing results, that ineffective practices are replaced with more promising ones, and that the money spent on
the programs is a sound investment” (p. 345). With the changes made in the legislation, more schools became eligible for Title I funding and rushed to implement schoolwide programs. What is known about effectiveness of these Title I programs at improving academic achievement of low-income students? Wong and Meyer (1998) stated that “the adoption of Title I schoolwide programs in schools and districts has been much faster than the availability of evaluation findings about their effectiveness” (p. 132). Title I was becoming a huge financial investment with little research to support the effectiveness of the programs. Funding for Title I had more than doubled from 6 billion dollars in 1992 to 12.3 billion dollars in 2004 (Clayton, 1991; Robelen, 2004). Wong and Meyer (1998) and Taylor and Teddlie (1999) stated that despite the huge financial investment of tax dollars in Title I and the rapid expansion of Title I programs in schools, little research regarding the nature of schoolwide programs and their effectiveness had been published. In light of the heightened demands of accountability and financial investment stemming from the federal government, there is a need to evaluate the effectiveness of Title I programs in schools thoroughly as they directly relate to student achievement of all identified subgroups, not just to the broad category of all students.

The metamorphosis of Title I was a predictable occurrence given the shifts in power in Congress from republican to democratic and back again. President Lyndon B. Johnson, a democrat, initiated the War on Poverty that led to the inception of Title I in 1965. In the early seventies, politicians sought to tighten regulations under the guidance of two republican Presidents, Nixon and Ford, by requiring districts to “buy in” to improving academic achievement by providing matching funds. During the late seventies, democratic President Carter relaxed the regulations to incorporate school-wide
programs. During the republican Reagan-Bush era of the eighties, there was a reduction in paperwork, a change of name from Title I to Chapter 1, and a removal of the matching funds requirement. The Republicans also expanded the school-wide initiative to provide more flexibility by funding class size reduction, increasing parent involvement, and improving teacher training. During the nineties, democratic President Clinton sought to tighten requirements once again by linking funding to improve the establishment of standards and the development of assessments. By the new millennium, republican President Bush added an additional layer of accountability with the AYP mandates and federal sanctions for failure to show improvement.

**Title I programs.** Since the inception of the ESEA in 1965 and the numerous of amendments to it, there have been a variety of ways in which districts have used Title I funds. Le Tendre and Chaibrán (1998) identified such common strategies districts use, including investing in curriculum and instructional materials, professional development, hiring specialists and master teachers, and using technology to support curriculum and instruction. Taylor and Teddle (1999) noted other districts included parental involvement as a part of their Title I plan. While Wong and Meyer (1998) reported that in some cases Title I funds were used to reduce pupil-teacher ratio. Sunderman (2001) reported similar results as she noted that schools used two strategies for Title I fund allocation: reducing class size and targeting specific students, subjects, and/or grades for remediation.

Many of the previously mentioned strategies required restructuring the existing school day without increasing the length of the school day. Often special classes were altered or eliminated to provide more instructional time for core areas such as writing,
mathematics, and reading (Sunderman, 2001). However, Deich, Wegener, and Wright (2002) reported that in approximately 40% of Title I schools some funding was used to support less intrusive programs that provided extended learning beyond the school day. In his study of Berkeley County School District in South Carolina, Etheridge (2001) reported that the district leaders used funds to implement extended-year programs for at-risk learners in Grades 3 through 8. Washington as summarized in Dodd and Wise (2002) reported that students who attended the Optional Extended Year Program in Austin, Texas showed measurable gains in their reading scores on the classroom assessments as well on the Texas Assessment of Academic Skills (TAAAS). Like the earlier pull-out programs, extended-day and extended-year programs allowed instructors to target specific students and/or specific content areas or grades. However, unlike the pull-out programs, the extended-day programs provided additional time (i.e., time on task) which minimized, if not eliminated, the disruption and curricular fragmentation caused by the pull-out programs.

Similar initiatives were being implemented in the primary, elementary, and middle and high schools in Riverview, a pseudonym for a diverse community nestled in Central Jersey. Riverview Public School District leaders continued to grapple with the ongoing issue of implementing effective programs that improved the performance of at-risk students and closed the achievement gaps. Like many other administrators, Riverview district leaders used Title I funds to hire classroom aides, fund pull-out programs, purchase instructional materials, reduce class size, and fund extended-day programs, yet very few documented studies of the effectiveness of these initiatives were conducted. The large percentage of African American and Hispanic students
participating in these programs and the increasing accountability mandates from the federal government led Riverview district leaders to initiate collection of empirical data on the effectiveness of their programs.

The Research Problem

Riverview leaders were not alone in their plight of poor performance of a substantial population of their African American and Hispanic students. Statistics from the U.S. Department of Education (2004b, 2004c) indicated that:

1. Hispanic children often don't attend school until they reach mandatory school age.

2. They [Hispanic students] have the highest dropout rates of any group in the country—more than 27% of Hispanic students drop out.

3. On the 2000 National Assessment of Educational Progress reading assessment, 40% of White fourth-graders scored at or above proficient, compared to only 16% of their Hispanic peers and 12% of their African American peers.

4. In math, Hispanic achievement also lagged: 35% of White fourth-graders scored at or above proficient. Just 10% of Hispanics and 5% of African Americans scored as high achievers.
5. Just 13% of Hispanic students get a college education.

(Hispanic Students and the Achievement Gap: African American Students and the Achievement Gap)

The data were of particular concern to district leaders because 23% of the approximately 1500 students attending the Riverview schools were from low-income families, and 49% of the total population is White, 16% African American, 15% Hispanic, and 20% Asian.

Concerned about the achievement of the minority students, Riverview district leaders began disaggregating their standardized testing data 5 years before the NCLB legislation made it mandatory. In response to their findings, leaders designed and implemented several programs aimed at improving academic achievement for all students by raising the bar and closing the gap. District leaders combined Title I resources with other district resources to fund and implement a variety of programs including class-size reduction (CSR), investment in research-based curriculum and instruction materials, professional development, hiring of aides, and creation of extended-day programs. Some of the early programs and initiatives failed and were no longer in operation. Although documentation regarding the reasons for failure did not exist, district leaders suspected that failure may have been due to implementation fidelity. Collins and Hanson as cited in Taylor and Teddlie (1999) “found that schools often established grandiose goals, but few components supporting those goals were actually implemented” (p. 303).

Riverview also had programs such as the extended-day programs which have continued to operate despite little documentation regarding the effectiveness of these programs. The accountability resulting from the NCLB legislation prompted Riverview district leaders to investigate not only the degree of implementation, but also the
effectiveness of all of these programs, particularly those Title I programs geared toward improving the performance of at-risk students.

Most information regarding the effectiveness of Riverview's programs came from informal feedback and staff perceptions. Teacher, parent, and student surveys were used to collect data regarding the perceived effectiveness of extended-day programs that primarily served low-income students. These extended-day programs provided reinforcement in the areas of mathematics and language arts literacy. Although data collected from parents, teachers, and students were used to modify the extended-day programs, no empirical studies examined the relationship between the Title I-funded extended-day programs and student achievement. Examination of the district results on the New Jersey state assessments for Grades 4, 8, and 11 revealed that not all subgroups showed improved performance over the last 5 years, and the achievement gaps still existed. Leaders questioned whether students who had participated in the extended-day programs over the last 5 years showed any significant academic gains from their participation. For years, district leaders had investigated, adopted, and facilitated the implementation of new programs, yet many leaders failed to incorporate the necessary feedback systems to provide empirical data on the effectiveness of the programs at obtaining the desired goals. As the federal government continued to use accountability measures such as NCLB and AYP to regulate distribution of funds, there was a growing need for today's administrators to be savvy researchers skilled in data collection and analysis, program assessment and evaluation, and data-driven decision making. Deich et. al. (2002) summed up the role of district leaders:
In order for initiative leaders to convince school administrators to allocate Title I funds to support extended learning activities, now more than ever, they must be able to clearly demonstrate the link between those activities and improvements in academic performance. (p.4)

However, evaluating the effectiveness of Title I programs was not an easy task. Over the past 40 years, the emphasis of Title I has changed from a source of new money that came with few restrictions, no plans for accountability, and no sanctions for school failure to a tightly regulated, under funded program with restrictions on fund appropriations, measures of accountability, incremental systems of mandatory sustained student improvement, and sanctions for failure to improve. Such “changing views about the government’s ability to affect poverty have made Title I’s effectiveness difficult to evaluate over time” (Cros, as cited in Cook, 2005, p. 24). Thus, the problem for this study could be expressed as a question: Was the investment in Riverview’s Title I-funded extended-day programs paying off?

The research and literature, discussed in more detail in chapter 2, provided conflicting evidence of the effectiveness of Title I programs. Research done by Etheridge (2001), Washington as summarized in Dodd and Wise (2002), and Borman and D’Agostino (1996) indicated that while students who participated in Title I programs showed some moderate academic gains, the overall effect at closing the achievement gap was negligible. However, prior studies did not differentiate the types of Title I programs thereby leaving room for speculation as to the effectiveness of one Title I program as compared to another.
Purpose of the Study

Recently the Riverview Board of Education commissioned an ad hoc committee to examine minority student achievement. Parents, teachers, administrators, students, and community leaders comprised the 60-member committee, and after meeting several months the committee found it difficult to agree on recommendations. Although the committee members agreed that minority student achievement was a concern, members were perplexed at finding a solution. They required data and unfortunately the district did not have them. Thousands of dollars had been invested in extended-day programs, and committee members wanted empirical evidence regarding the effectiveness of these programs. There were no discernible feedback systems established, few written documentations of earlier programs, and with frequent district administrative turnover, much of the information was simply lost. Significant to the researcher was a question asked by a parent attending a board of education meeting: “How do we know these programs are working?”

The purpose for this study was to determine if there was any measurable relationship between mathematics achievement of second through fifth-grade students and participation in Riverview’s Title I-funded extended-day programs. This program evaluation included data from the 2004–2005 school year and should be seen as a pilot study for initiating continuous evaluation for improvement. In this investigation the researcher examined the mathematics achievement of Title I-eligible students who participated in the extended-day programs as compared to: (a) the mathematics achievement of all non-participants, (b) the mathematics achievement of Title I-eligible students who did not participate in the extended-day programs, and (c) district, state, and
national test norms. In a second investigation the researcher compared the mathematics
achievement of: (a) non-Asian minority participants with that of White non-participants
and (b) lower SES students with that of moderate/higher SES students. Third the
researcher sought to examine the nature of existing achievement gaps in the Riverview
Public Schools by comparing the pretest scores of extended-day participants with those of
non-participants.

The researcher aimed to add to the current body of research, which is particularly
limited in the area of the influence of Title I-funded extended-day programs on academic
achievement. The researcher sought to provide practical implications for district leaders
by not only examining the effectiveness of the extended-day programs, but also the key
components of the programs that contributed to them being effective or ineffective. With
NCLB legislation, the federal government has ultimately placed the high stakes
accountability in the hands of district leaders who must make decisions regarding
curricular programming. The researcher sought to provide district leaders a platform on
which to evaluate extended-day programs designed for Title I-eligible students, make
informed choices regarding the cost effectiveness of such programs, and conduct future
research regarding the sustained effects of participation.

**Delimitations and Limitations**

Although the present study served to add to the current research on the
relationship between participating in extended-day programs and mathematics
achievement, district leaders must use caution when making generalizations based on the
results of this study as several limitations and delimitations apply. The researcher imposed the following delimitations:

1. This researcher examined students in Grades 2 through 5 attending the Riverview School District. The researcher did not investigate students in other grades or school districts.

2. The researcher only examined extended-day programs funded by Title I. The researcher did not investigate other Title I schoolwide programs such as class size reduction, the hiring of master teachers, the use of classroom aides, or pull-out programs for remediation, although these variables were discussed in the study.

3. Students who participated in the gifted and talented (G&T) and special education math programs were excluded from the study due to the differences in their schedules.

4. The researcher used data collected from the regular testing using NJ ASK 3 and 4 tests, the Terra Nova tests for Grades 2 and 5, and internal District Benchmark Assessments (DBAs).

5. Random sampling was not used. To the extent possible, the researcher used matched groups to control for extraneous variables.

The researcher also noted the following limitations of the study:

1. Selection of participants in the extended-day program was determined by the building principals and teacher committees, and participation in the extended-day programs, although highly encouraged, was by voluntary parental permission. This affected the randomization and the overall sample size.
2. The internal DBAs were created by Creative Learning Corporation and modified by teachers which may have affected the reliability and/or validity of the instrument.

3. The internal district assessments were administered and graded by classroom teachers which may have introduced a level of bias.

4. Since randomization was not possible, substantial pretreatment/pretest differences may have existed among the three groups of students. In measuring pretest-posttest scores, a regression toward the mean posed a potential problem, as students in the lowest pretest percentile ranks had a greater chance of earning more gain points than did students in higher percentile ranks.

5. Student attendance at all sessions was encouraged yet could not be controlled. Thus disparities in time on task may have affected the outcome.

6. No information was available regarding student participation in other achievement-enhancement programs outside the school day.

7. No information was available on amount of time parents spent helping their children prepare for benchmark assessments and standardized tests used as instruments in this study.

8. Although teachers recorded the activities and learning goals on a daily basis, there were no formal observations of the program conducted by the researcher and there were subsequently no observable assurances that instructional strategies were implemented correctly, the curriculum was incorporated properly or that time on task was maximized.
Despite the limitations and delimitations, the present study had potential to add to the current body of research on Title I-funded program implementation fidelity and the effectiveness of these programs.

Research Questions

Title I was instituted to increase the academic achievement of low-income students and close the achievement gap between these students and their more economically stable peers. To measure any positive gains at closing the minority and SES achievement gaps, one must first establish that these gaps exist. Therefore the first task was for the researcher to explore if there were preexisting differences in mathematical ability between the various populations in the present study. The first research question was stated as: (1) What is the significance, if any, of preexisting differences in mathematical abilities of the participating groups in this study? The answer to this research question helped establish a baseline for comparison of participating and non-participating students.

The researcher investigated the relationship between participation in extended-day programs and the mathematics achievement of Title I-eligible students. One premise behind Title I was that participation in Title I programs should lead to an increase in mathematics achievement. This hypothesis was the basis for the second relevant research question: (2) What is the relationship between participation in Title I programs, specifically the Project Excel program, and mathematics achievement? Clark as summarized in Huang (2001) stated that “students who spend 20 to 35 hours on
constructive learning activities after school do better in regular school than students who
do not have these after school experiences (p. 35).

A third research question arose to investigate the influence of participation in
Project Excel at closing the achievement gaps. The question can be stated as: (3) What is
the relationship between participation in Title I programs, specifically the Project Excel
program, and closing the achievement gaps between low-income and moderate/high-
income students and non-Asian minority and White students? This research question can
be interpreted on two levels. Borman and D’Agostino (1996) summarised this dilemma:

The goal of ‘narrowing the achievement gap’ between at-risk children and
their more advantaged peers has two distinct definitions that have
influenced the selection of an appropriate control group. Some
researchers have responded to the question: does Title I narrow the
achievement difference between program participants and a nationally
representative sample of non-participants? Alternately, others have
questioned whether this gap would widen without the existence of Title I
services. Researchers have responded to the former question by
comparing Title I students to all other non-participating students, while the
latter question has been addressed by comparing Title I students to
similarly needy controls who did not receive compensatory education,
services. (p. 311)

The researchers investigated the achievement gap from both vantage points.

Simply participating in Title I programs was not a guarantee of success. In their
review of extended-day programs Dodd and Wise (2002) stated that “It is clear that
increasing the school day without quality instruction will not have a profound effect on student achievement... A higher success rate is found when students participate in extended-day programs that are similar to their regular school-day programs” (p.1). With regard to contribution of other factors, such as time on task and small class size, the researcher included a discussion on these additional variables but did not evaluate them separately. Le Tendre and Chalarín (1998), Edwards, Katz and Brenton (2001), Heifston and Wright (2003), and Huang (2001) suggested that factors such as time, instructional materials, staff selection and training, instructional methodologies, class size, and standards and expectations of staff were relevant factors that influenced the success of extended-day programs.

Definition of Terms

Riverview district leaders incorporated the use of annual standardized tests and District Benchmark Assessments (DBAs) to analyze and evaluate the performance of students participating in the Title I-funded extended-day programs. District leaders sought empirical data as to the effectiveness of these programs in improving academic achievement of poor and low-income students. Collected data were used to make informed decisions regarding continued funding and program modifications and adjustments. A list of defined terms is included below for the purpose of this study: Adequate Yearly Progress (AYP) - The minimum standard for percentage of students identified as proficient for each year during that period. For purposes of this study the researcher limits the application of AYP to students taking the NJ ASK 4 mathematics test. The AYP for the 4th grade math assessment increases incrementally over a 12-year

District Benchmark Assessments (DBAs) – DBAs represent a series of mathematics assessments given to students in grades PK through 5. The DBAs consist of a pretest given in September, a midyear assessment given early February, a posttest given in June and nine monthly assessments. The purpose of these assessments is to provide the classroom teacher with information regarding students' level of proficiency in mathematics. The DBAs were designed by Everyday Learning Corporation as a part of the Everyday Mathematics program. DBAs were edited by classroom teachers.

Extended-day Programs – These are programs that provide an extended time for learning that goes beyond the school day. The programs may come in a variety of forms—including before and after school programs, extended-day kindergarten, Saturday programs, summer programs, and programs held during holiday and vacation time when school is closed (Dodd & Wise, 2002, p.1).

Low-income – The federal government defines low-income as a family of four with a combined household income between $18,850 and $37,780 (U.S. Department of Education, 2004a).

National Assessment of Educational Progress (NAEP) – NAEP is often called the "Nation's Report Card." It is the only measure of student achievement in the United States where you can compare the performance of students in your state with the performance of students across the nation or in other states. NAEP, sponsored by the U.S. Department of Education, has been conducted for over 30 years. The results are widely
reported by the national and local media (National Center for Education Statistics NCES, 2005a).

**New Jersey Assessment of Skills and Knowledge (NJ ASK)** - The NJ ASK is a criterion-referenced test that assesses student achievement in the knowledge and critical thinking skills defined by the New Jersey Core Curriculum Content Standards in language arts literacy, mathematics, and science. The test is currently given to third- and fourth-grade students (NJ Department of Education, 2004a).

**New Jersey Grade Eight Proficiency Assessment (GEPA)** - The NJ GEPA is a criterion-referenced test that assesses student achievement in the knowledge and critical thinking skills defined by the New Jersey Core Curriculum Content Standards in language arts literacy, mathematics, and science. The test is currently given to eighth-grade students (NJ Department of Education, 2004b).

**New Jersey High School Proficiency Assessment (HSPA)** - The NJ HSPA is a criterion-referenced test that assesses student achievement in the knowledge and critical thinking skills defined by the New Jersey Core Curriculum Content Standards in language arts literacy, mathematics, and science. The test is currently given to 11th-grade students (NJ Department of Education, 2004c).

**Poor** - The federal government defines poor as a family of four with a combined household income less than $18,850 (U.S. Department of Education, 2004a).

**TerraNova Test** - Norm-referenced test given to second- and fifth-grade students during the spring. The test assesses mathematics and language arts. *TerraNova* CTBS Multiple Assessments editions, for Grades 1 through 12, combine selected-response items with
constructed-response items that allow students to produce short and extended responses (CTB/McGraw-Hill, 2004).

Trends in International Mathematics and Science Study (TIMSS) - TIMSS was developed by the International Association for the Evaluation of Educational Achievement (IEA) to measure trends in students' mathematics and science achievement. Offered in 1995, 1999, and 2003, TIMSS provides participating countries with an unprecedented opportunity to measure students' progress in mathematics and science achievement on a regular 4-year cycle. The next cycle of TIMSS is scheduled for 2007 (NCES, 2005b).

Summary of Chapter One and Organization of Remainder of Study

In chapter 1, the researcher examined the history of Title I, the regulatory changes that have occurred since its introduction, and the various programs that fall under the umbrella of Title I-funded programming. Many district leaders were seeking solutions to the rising problems of inequitable and disproportionate student achievement. Since the inception of Title I, district leaders have employed many projects including extended-day programs to improve academic achievement of low-income students. Many of these programs failed to achieve their goals, yet little empirical data existed for detailed analysis. In the present study, the researcher sought to collect empirical data to determine the relationship between participation in Title I-funded extended-day programs and mathematics achievement.

A review of research and literature is presented in Chapter Two. This section includes a review of research that suggests that the participation in Title I programs
improves the academic achievement of low-income students, and contradictory literature that suggests that students who participate in Title I programs show no measurable, positive differences in academic achievement from those who do not participate. Chapter 2 also includes studies in which researchers examined other factors affecting student achievement including class size, time on task, instructional methodologies, staffing, professional development, curriculum, and program evaluation which represent important components that make up the theoretical framework of the study.

Chapter 3 provides a description of the research design, participants, and instrumentation used in the present study. This section includes details of the type and structure of the extended-day program, strategies employed to answer the research questions, the method for data collection and analysis, hypotheses tested, and variables examined in the present study.

The results of the investigation are presented in chapter 4. A detailed statistical analysis of the data and an interpretation of the descriptive and inferential findings that link to the research questions are the focus of this chapter.

The pertinent findings of the investigation and their relevancy are discussed in chapter 5. Connections are made between prior research, current findings, and considerations for policy, practice, and future research.
Chapter II.

Literature and Research Review

Introduction

"Despite [the] seeming wealth of [Title I evaluation] information, the overall educational effectiveness of the program has remained a matter of debate" (Borman & D'Agostino, 1996, p. 309). Much of the debate was centered on the inconclusiveness of the results of program effectiveness. However, two mitigating factors also affected the outcome of the evaluation: (a) Did the research designs of prior studies lend themselves to sound conclusions regarding program effectiveness and (b) Did the research adequately account for the impact other variables that influence program effectiveness?

In a study of New York City schools, van der Klaauw (2005) described the shortcomings of prior research on Title I program effectiveness. He found that much of the earlier research on Title I compared the pre- and posttest performance of Title I students to a norm-referenced group. Normal Curve Equivalents (NCEs) were used to measure the gain scores obtained. Such a model is slightly flawed in that it "assumes that without program services, the student would stand at the same level on the post test as the pre-test, which would be indicated by a zero change score...This is commonly referred to at the equipercentile assumption" (van der Klaauw, 2005, p. 10). For this model to be useful, the norm-reference group would have to be nearly identical to the Title I group.

A similar problem existed with research methods that compared the academic performance of Title I students with that of non-Title I students (van der Klaauw, 2005).
Even when examining gain scores, the amount of pre- and posttest mobility was not inherently equal. A non-Title I student from a moderate income may have a high level of motivation and skills and thereby produce significant gains, whereas a Title I student receiving the same regular instruction plus additional supplemental instruction may show gains as a result of the additional instruction. However, these gains may not be significantly higher than or even equal to the gains made by the non-Title I student.

A third concern van der Klaauw (2005) noted was that most of the prior research hypothesized whether Title I has been successful at closing the achievement gap, rather than the more pertinent issue of whether it has been useful in preventing Title I students from falling further behind. Would the gap be wider if Title I did not exist over the last 40 years? Proponents of Title I would argue that it was never intended to maintain the gap, as suggested by van der Klaauw, but rather reduce it. If the initial goal was to reduce the gap then the program's level of effectiveness should be measured according to the degree in which the goal was achieved. The present research served to address both issues by comparing the academic performance on three different levels: (a) Title I-eligible participants compared to Title I-eligible non-participants, (b) Title I-eligible participants compared to Title I-eligible non-participants, and (c) Title I-eligible participants compared to district, state, and national norms.

Evaluating the effectiveness of Title I programs has also been complicated by the variety of Title I programs and other contributing factors that may influence student achievement. Title I funding has supported many different programs and initiatives. Often, programs overlapped in their services and a single child might have participated in an after-school program, tutoring, a basic skills program, and pull-out programs all
designed to increase mathematical achievement. Which program or combination of programs was responsible for the achievement gains, if any? It was difficult to discern without conducting a thorough investigation.

A further complication was the existence of other variables that have been shown to affect student achievement such as the mathematics class size, the amount of daily time spent learning mathematics, the level of parental involvement, the level of experience of the teacher, the degree in which the teacher implemented research-based methods of mathematical instruction, the types of mathematical instructional materials, and the degree of high expectations and beliefs that students will succeed held by teachers. In this study the researcher considered some of these variables while examining the influence of participation in one type of Title I program, an extended-day program, on the overall mathematics achievement of Title I-eligible students.

Historical Reviews

The Equality of Educational Opportunity (EEO) study conducted by Coleman et al. in 1964 was one of the earliest and most notable research studies conducted in the field of education (Haller & Klein, 2001). What began as "an obscure provision of the Civil Rights Act of 1964...to ascertain the 'lack of availability of equal educational opportunities'..." (Haller & Klein, 2001, p. 22) resulted in a landmark study that would spawn future decades of research, educational endeavors, and reforms. Americans were not shocked that the findings suggested that African American students achieved at lower levels than did White students; however, it was surprising to discover that these students often achieved at levels more than one standard deviation below their White counterparts.
Of particular interest from the EEO report was the relationship between socioeconomic status (SES) and student achievement. Of the factors measured by the EEO, SES, responsible for only approximately 10% of the variance in students' achievement was the most powerful determinant of academic success (Haller & Kleine, 2001). Despite the outcome of the report, "America has pinned its hopes on the public school as a means to offset at least some of the advantages that accrue to the children on the more fortunate" (Haller & Kleine, 2001, p. 37) and thus the birth of Title I.

During the 60s and 70s few studies actually examined data collected from Title I students and compared those data to other students. The Sustaining Effects Study conducted from 1976 to 1979 represented one of the earliest studies that actually did (Carter, 1984). The findings of the study indicated that students participating in Title I programs had higher achievement gains than did students from low-income families who did not participate in the programs. However, Title I students did not achieve at the same level as students who were not from low-income families. The interpretation of the research led to two different conclusions: (a) Title I programs were effective, because there was evidence of student achievement gains and (b) Title I programs were ineffective, because they failed to close the achievement gap as intended (Borman & D'Agostino, 1996). Since the Sustaining Effects Study many researchers have investigated both conclusions.

Efforts to Improve Academic Achievement

After 40 years of existence, evidence of the relationship of Title I programs to increasing academic achievement of low-income students is still inconclusive. After their
review of the Chapter 1 Instructional Program Report by Neitzworth, Cisneros, and Sanchez. Wong and Meyer (1998) reported that few significant differences were found between Houston schools that incorporated schoolwide programs and those that did not. For the few significant differences that were found, the differences were small. In short, the reading and math scores were significantly higher in Grades 2 and 4 in Houston schools that did not incorporate schoolwide programs. Similar results were found when they examined limited-English-proficiency (LEP) students whose math scores were slightly higher in the non-schoolwide program schools. Their results not only suggested that schoolwide programs were not helpful, but may have also implied that the presence of schoolwide programs may actually cause a decrease in achievement.

In a study of New York City schools, van der Klaauw (2005) noted a similar result. He used school data collected over a 7-year period and a regression-discontinuity approach to compare the academic achievement of students in Title I schools and non-Title I schools both with cutoffs on close to the poverty rate. “Overall, the estimates show little evidence that Title I funding leads to improved test results. In fact, most of the effect estimates imply a negative effect on performance which in several cases is statistically significant” (van der Klaauw, 2005, p. 19). Evidence from 1993, 1997, and 2001 New York City school data suggested that there was no statistical evidence to support the claim that students performed better as a result of Title I funding; yet there was statistical evidence of the negative effect of participation in Title I programs.

On the contrary, Wong and Meyer’s (1998) review of the work done by Davidoff and Pierson stated that in Philadelphia schools which incorporated schoolwide programs students met more state-established criteria for reading and math than did students from
schools that did not incorporate schoolwide Title I programs. When they compared all four measures of student achievement for both schoolwide and non-schoolwide programs, they stated that a "higher percentage of schoolwide program sites met each of the four state-established criteria for reading achievement than did non-schoolwide program sites. In mathematics the percentage of schoolwide program sites was higher for only three out of the four criteria" (Wong & Meyer, 1998, p. 133).

Etheridge (2001) researched the effectiveness of schoolwide programs in Berkeley County School District in South Carolina. He examined the math and reading scores of students in fourth and seventh grades. Student achievement was measured by a reduction in the percentage of students in Q1 (scoring below the 25th percentile). He noted that in a 2-year period, Title I students had experienced a decrease in the percentages of student in Q1 with the most significant decrease being a 19.6% drop for seventh-grade mathematics. His study showed an increase in percentages of Title I students scoring above the 50th percentile for all groups except for seventh-grade reading.

Taylor and Toddle (1999) investigated the implementation fidelity of Title I schoolwide programs at 10 urban elementary schools. Examinations of both norm-referenced and criterion-referenced tests results "do not reflect implementation of schoolwide programs that improved the instructional program" (p. 307). They proposed that one factor contributing to the test results was the limited extent to which the schoolwide programs were implemented as proposed (fidelity).

Borman and D'Agostino's (1996) research using data derived from 17 federal studies conducted between 1966 and 1993 which examined the reading and math scores of students in Grades 3 through 12, yielded positive support for Title I programs. Using
regression analysis of weighted effect sizes on mediating variables, they were able to isolate some positive gains in test scores associated with participation in Title I programs. However, they were not able to make positive statements about the robustness of the findings. They summarized their findings as:

Contrary to widely held beliefs regarding the historical stability of programmatic impact, the results suggest a positive trend for the educational effectiveness of Title I across the years of its operation. Although the Title I population effect estimate cannot be determined from existing federal data from a summative perspective, the Title I program appears to have contributed to the achievement growth of the children it served. (p. 324, emphasis added)

Using a simple t test, Edwards, Kahn and Breton (2001) examined the pre- and posttest gains of students participating in the Math Corps Summer Camp. "The data consistently shows (sic) statistically significant mean increases from pre- to posttest, ranging from 26 to 60 percentage points (p < 0.001 in each case)" (p. 422). Although the researchers were careful not to form any strong conclusions because of their lack of a control group and their poor design, the data -- when taken in conjunction with other studies -- supported the premise that extended time on task may lead to increased achievement.

Some researchers found that extended-day programs did increase student achievement during the school year, however, lack of continued participation over the summer led to dramatic losses of achievement gains. Boram and D'Agostino (1996) noted that students participating in Title I programs did show more academic
achievement when compared to similar peers who did not participate in the programs. However, they noted that when the testing cycle was fall-spring, student achievement gains were higher than when the testing cycle was spring-fall or annual. Their review of David and Pelavin’s research noted that students participating in Title I programs did show large achievement gains; however, the “demonstrated large achievement gains by Title I students over the school year typically [were] followed by diminished summer growth or achievement losses” (p. 311). They attributed this difference to the summer effect, which was particularly more deleterious to at-risk students in intermediate and upper grades than to at-risk students in primary grades. Borman and D’Agostino proposed that:

[The substantial fall/spring math gains that young children in the early elementary grades achieve are more prone than reading to the negative consequences of the summer effect. Therefore, summer math programs for students in the early elementary grades may counteract the summer effect and sustain their rapid growth beyond the school year. (p. 324)]

Other researchers also found that although the extended-day programs have shown some moderate increase in academic achievement, the gains were lost over the summer. Borman (2003) found, “[s]trong research evidence suggest that the achievement gap is already large when children enter kindergarten, that the gap increases as students go through school and that summer learning differences may play an important role” (p. 50). Cooper et al. as summarized in Borman (2003) estimated that the average child loses over one month’s worth of knowledge in math and reading during the summer break. Middle/upper income children were more likely to have summer experiences to
counter the summer loss than did low-income children. Entwisle and Alexander as cited in Edwards, Kahn, and Brenton (2001) "studied the deleterious effects of summer vacations on the academic progress of low-income African American and White children through the first and second-grades" (p. 421). They found that poor children of both races suffered from an academic summer setback. They also noted that summer school programs did little to eliminate that effect. They hypothesized possible explanations for such poor results were the casual nature of summer programs and the possible stigmatization of students participating in remedial programs.

**Efforts to Close Achievement Gaps**

Research on the influence of Title I programs at improving student achievement has been inconclusive, and similarly the research regarding Title I programs effect at closing the achievement gaps was also inconclusive. In his review of Title I research, Pitsch (1995) reported that "the research has shown that most Title I programs do not boost the performance of disadvantaged students enough to close the achievement gap between those low achievers and their more advanced peers" (p. 1). Yet Wong and Meyer's (1998) evaluation of the research done by Wong, Sunderman and Lee showed that the achievement gap between Title I-eligible and ineligible students was significantly smaller in Minneapolis schools which incorporated schoolwide programs, than in similar schools which did not incorporate these programs. The same study also showed a significantly smaller gap in Houston schools which did not incorporate schoolwide programs.
Etheridge's (2001) study produced some apparent shrinkage of the achievement gap. Etheridge reviewed the reduction of the percentage of students scoring below the 25th percentile in Berkeley County School District. He reported that "[a]lthough the percentages of non-Title I students scoring at the Q1 level [below 25th percentile] continued to be less than those of Title I students, improvement in performances, as indicated by the 'percentage point change' and 'gap reduction' were encouraging" (p.338). Although small increments of gap shrinkage were apparent, none of the research overwhelmingly supported the idea that Title I programs closed the achievement gaps.

Perhaps a flaw exists in the idea that the gaps can be closed; for a gap to be closed, one of two things must occur: (a) The underprivileged children must learn at a faster rate than their privileged counterparts until they eventually catch up or (b) the privileged students learning rate must remain stagnant or decrease as the underprivileged students catch up. Each of the previous theories assumes no regression on the part of the underprivileged during the summer break. A third possibility to eliminate the gaps is to prevent the gaps from occurring by addressing the educational needs of children prior to beginning formal schooling.

Other Related Factors

Theoretical framework. One cannot establish a strong relationship between extended-day programs and student achievement without considering other factors that affect student achievement. Such variables such as extended time, researched-based instructional methodology, class size, staffing and staff training, curriculum, and degree of program implementation (fidelity), monitoring, and evaluation have been shown to
impact student achievement and make up the theoretical framework for this study. In developing their extended-day programs, Riverview district leaders accounted for and incorporated by design provisions for extended time-on-task, small class sizes, highly qualified teachers, staff training, researched-based methodologies, and instructional programs and program evaluation. They anticipated that the outcome of the extended-day programs would be improved student achievement and closing of the achievement gaps. Figure 1 depicts the theoretical framework for this study. A brief discussion as to how the factors outlined in the framework are related to extended-day programs follows.

Time. With the demand for higher standards and the impact of accountability systems placed on states and districts, educators were striving for more creative ways to increase student achievement. As more and more was added to the daily curriculum a child experienced, the length of the school day has remained relatively stagnant. Even the current system of education where students have the summer off is reflective of the primitive agrarian society we once had. In Prisoners of Time, the National Education Commission on Time and Learning (1994) stated:

If experience, research, and common sense teach nothing else, they confirm the truism that people learn at different rates, and in different ways with different subjects. But we have put the cart before the horse: our schools and the people involved with them—students, parents, teachers, administrators, and staff—are captives of clock and calendar. The boundaries of student growth are defined by schedules for bells, buses, and vacations instead of standards for students and learning. (PowerPoint slide 3)
Figure 1. Theoretical framework for the present study.
Even the courtrooms have become home for extended time litigation issues that recognize the need to provide additional support for at-risk students. “At the trial of CFE v. State, numerous expert witnesses confirmed the widespread consensus that providing additional time on task is an essential part of ensuring that at-risk students have the opportunity for a sound basic education” (Campaign for Fiscal Equity, Inc., 2002, p.1).

One of the earlier advocates for extended time was Bloom (Berliner, 1990). When referring to the nature of human learning, Berliner (1990) summarized Bloom’s theory of mastery learning as an endeavor in pacing. All students can learn if given sufficient time to learn. Bloom suggested that time be open-ended with reteach cycles, and students should persist until they meet the criterion. Much of Bloom’s learning theories have been substantiated over the years. In their review of the research on time, Dodd and Wise (2002) reported that some students required three to six times more time to learn than did the average student. “It is now clear that more time is necessary to support those students in need of additional time to be able to successfully master and build upon national and state standards and expectations” (Dodd & Wise, 2002, p.1).

In 1983, the National Commission on Excellence in Education (NCEE), published A Nation at Risk, in which they called for “more effective use of the existing school day, a longer school day or a lengthened school year” (NCEE, Recommendation C).

“Although proposals to extend the school year were considered in 27 states…very few were actually approved. Moreover, none of the states passing legislation during that time increased the school day beyond 6½ hours or the school year beyond 180 days” (WestEd, 1998).
A decade after *A Nation at Risk*, not much had changed as shown by the National Education Commission on Time and Learning (1994) as summarized in Dodd and Wise (2002) who reported that “American public schools have held time constant and let learning vary” (p.1). Ten years after the National Education Commission on Time and Learning, Deich, et al. (2002) in reviewing the NCLB legislation, summarized that “[t]he new bill calls for increases in the amount and quality of learning time, through extended school-year, before- and after-school programs, summer programs, and other opportunities that enhance the school-day curriculum” (p.4). Now with increased accountability more districts were turning to extended-day programs as a means of increasing academic achievement.

In a summary of research, Cotton (1989) noted substantial support for a positive relationship between time-on-task and student achievement. Cotton summarized:

Since student performance is dependent upon the amount of time

NEEDED to learn as well as the amount of time provided, only the

students who need greater amount of time to learn (and who, perhaps, do

not normally have enough time to pursue tasks thoroughly and learn they

well) perform better when they are given and make use of additional

learning time. (p. 8)

Several districts that offered extended-day programs have seen some success. Etheridge (2001) researched Berkeley County Public Schools which offered an extended-year program for at-risk students in Grades 3 through 8. The program ran 17 days during the summer and was a 6-hour day that included lunch, parent conferences, computer-aided instruction assistants, and instructional materials. Following its initiatives the
Berkeley County district leaders had seen a decrease in the number of third- and fourth-grade students performing below level on the state assessment.

Decisions regarding time for remediation formally made by school leaders were now being made by state policy makers. Under the NCLB legislation, schools and districts not meeting the established AYP criteria in a given area for 3 consecutive years were required to seek supplemental educational services. These services were not meant to supplant but to supplement the regular instruction in math and language arts. Many of these services came in the form of tutoring offered before or after school and on weekends. This mandatory allotment of time and resources was designed to help improve the academic achievement of students in a given population who were at-risk for not meeting proficiency standards. School and district leaders were held accountable to show that extended time led to enhanced performance.

Although these studies indicated that students benefited from additional time, a distinction of the type of time must be made. The research on Improving Student Achievement by Extending School (WestEd, 1998) and studies conducted by Berliner (1990) defined time as: allocated time, engaged time and academic learning time.

Allocated time referred to the total number of days or hours students were required to attend school. This included time spent in class as well as time spent during lunch and recess. Engaged time was defined as instructional time allotted for teaching. This included time spent not only on teaching, but also on things that occurred during teaching time such as taking attendance, handling discipline issues, responding to interruptions and announcements. Academic learning time was defined as the period when “the instructional activity is perfectly aligned with a student’s readiness and learning occurs”
The goal of the Project Excel program was to provide additional individualized instruction that maximized the potential of aligning activities with student readiness. This was accomplished by extending the school day and providing 60 to 75 minutes of additional mathematics instruction weekly.

Peterson (2005), director of the Afterschool and Community Learning National Resource Network stated “[w]ell-structured after-school programs effectively expand learning time for students,” (p. 11) but were these extended-day programs the most cost-effective use of Title I funds? WestEd (1998) summarized a study conducted by the Institute for Research on Educational Finance and Governance (IREFG) which examined the influence of four variables—time, peer tutoring, class-size reduction, and computer-assisted instruction on student achievement. Although the IREFG erroneously classified their research on pupil-teacher-ratio as class-size reduction, they stated that “increasing time was the least cost-effective of the four interventions in terms of math performance” (WestEd, 1998). The dollar bottom line has been the incentive to cut many programs, including those showing promise, yet when it came to assessing cost-effectiveness of a program many district leaders were ill-equipped with the necessary data to make that decision. How much time should be allotted to decide if an intervention program is successful and cost-effective...1, 2 or more years? In the face of fiscal restraints, Riverview district leaders wrestled with the question of whether Project Excel was cost-effective and they anticipated that this study would provide some guidance.

*Instructional practices.* Extending the school day without attention to quality of instruction may be a waste of time and money. Extended-day programs must incorporate research-based instructional methodologies to be successful. Huang (2001) stated “that
effective afterschool programs provided opportunities for students to acquire new skills and broaden their education through enrichment activities that complement their regular school day" (p.56). Muijs and Reynolds' (2000) study of primary schools in the United Kingdom identified several instructional strategies pertinent to teaching mathematics successfully. Such factors include linking new knowledge with prior knowledge, asking a lot of questions that involve pupil construction of knowledge, providing praise, using a variety of teaching strategies and materials, and creating a climate of learning using a constructivist approach. Muijs and Reynolds (2000) stated that:

We agree with Brophy's (1986) contention that effective teaching is likely to be a conglomerate of behaviours. It is unlikely that one isolated behavior will make the difference. Rather, it is the combination of effective teaching behaviors that will lead to better performance in pupils. The extended-day programs in this study strived to incorporate these research-based strategies. (p. 278)

Inherent in the design of the Project Excel program was the implementation of various instructional strategies. Operating under the premise of individualized instruction, teachers were trained to match their instructional strategies to the learning styles of their students. In many cases the Project Excel students were taught by their regular math teacher who knew their strengths and weaknesses as well as their learning styles. These teachers were also quite familiar with the researched-based pedagogy that supported the implementation of the Everyday Math program and the constructivist approach to teaching.
Parker and Gerber (2000) studied the effects of a science intervention program on middle-grade student achievement and attitudes. Eleven African American students who were identified as economically disadvantaged and whose mathematics and reading performances during the school year were below average, participated in the 5-week summer program. The focus of the study was on the influence of providing additional learning time and using constructivist instructional strategies. The program used a hands-on, inquiry-based approach to teaching science which was supported by research (Parker & Gerber, 2000). A correlational t-test performed on pre- and posttest criterion-referenced tests yielded a t-score of 5.52 (p < 0.001). This finding paralleled Mujsa and Reynolds' (2000) support of Brophy's contention that it is a combination of teaching factors that supports academic achievement.

Class size. One of the tenants of NCLB was the provision for funds to be allocated for class size reductions -- not be confused with pupil-teacher-ratios. Consider the following example to illustrate the distinction. An inclusive class of 24 children is co-taught by a certified math teacher and a certified special education teacher on a daily basis. The class size remains 24 as each day each teacher must teach 24 children. However, the pupil-teacher-ratio for that class is 12:1. Pupil-teacher-ratios further exacerbate true class size as they represent the “number of students in a school or district compared to the number of teaching professionals” (Lewis & Baker, as cited in Achilles, 2005, p. 12).

Project STAR (Student Teacher Achievement Ratio), a longitudinal investigation of class size in Tennessee, was most noted for providing evidence of the effectiveness of class size reduction on student achievement. This randomized study tracked the
performance of some 1,600 students who participated in one of three groups: (a) small class size (μ = 15), (b) regular class size (μ = 25), and (c) regular class size with an aide (μ = 25), and who had test results over a 4-year period. Participation in small classes for consecutive years in the earlier grades had several lasting effects. The results from the STAR study “leave no doubt that small classes have an advantage over larger classes in student performance in the early primary grades” (Achilles & Finn, 2001, p. 10). Finn (2006) noted some of the lasting benefits:

At the extreme, for all students combined, the effects of attending small classes for four years increased the odds of graduation by 80%. For students from low-income homes, 3 years of small classes increased the odds of graduating by approximately 67%, and 4 years in small classes more than doubled the odds. The graduation rates for low-income students with 3 or more years of small class participation were at least as high as those of higher-income students, closing the gap in graduation rates completely. (p. 2)

Since the Tennessee STAR report, other researchers have sought to determine the lasting effects sustained by STAR participants. Finn (2006) summarized the a follow-up study conducted by Krueger and Whitmore (2001) who noted that students who participated in small class sizes were more likely to take college-entrance exams than were students who did not participate in small classes in Grades K through 3. Krueger and Whitmore (as cited in Finn, 2006) reported “the benefit for Black students was substantially greater than for [White] students, reducing the Black-White gap in college-entrance-test taking by 54%” (p. 2). Finn, Fox, McClellan, Achilles and Boyd-Zaharias as cited in Finn (2006) stated “small-class participation had a significant
positive impact on the amount of foreign language taken, and the highest levels taken in foreign languages and mathematics” (p. 3).

There have been several additional studies that have supported the benefits of small class size. The June 1998 Educational Issues and Policy Brief produced by the American Federation of Teachers, highlighted several of these studies. Rouse’s study of Milwaukee P-5 schools as cited in the AFT Policy Brief stated that “[s]tudents in the P-5 (small class size) public schools made faster math gains than students in regular public schools and the public magnet schools, and the same math gains as the voucher schools” (p.1). Maier and colleagues conducted a study of Wisconsin schools commonly known as the SAGE program. Maier noted that African American males participating in SAGE saw a rise in test scores of 56 points compared to 39.4 for the control schools (AFT, 1998). Etheridge’s (2001) study of Berkeley County showed that students in reduced class sizes (15 or fewer students) showed a 7 to 11% increase in math and language arts tests scores over the scores of students in non-reduced class sizes (18 or more students). In their review of Mac Iver’s research, Dodd and Wise (2002) reported that “providing content and instructional pace adaptations to accommodate the student’s style of learning during the extended learning time in one-on-one or two-on-one tutorial sessions can cause a rise in student achievement scores” (p.1).

The ultimate small class size is one-to-one, yet such individualized instruction would be expensive and impractical in a public school setting. However, the lessons learned from smaller classes should provide a stable foundation for the design and development of extended-day programs. Riverview district leaders anticipated that
Students in the Project Excel program, most of whom were minority students, would benefit from a small class size of four or five students.

**Staffing and Staff Training.** Thus far, the literature review has considered the effects of time, instructional practices, and class size on student achievement. What is the relationship between teacher training and student achievement? The federal government believed that there was a positive correlation between professional development (PD) and student achievement. “Section 1001 of the NCLB recommended increased professional development to elevate the quality of instruction delivered to students” (Tienken & Achilles, 2003, p. 153). Ironically, NCLB required districts to select and to prove the use of scientifically based research (SBR) when implementing programs funded by Title resources, yet SBR supporting a profound positive relationship between PD and improved student achievement was lacking. Achilles (2002) examined over eight studies that supported professional development and confirmed all eight failed to find empirical evidence of improved student performance or teacher change. Tienken and Achilles (2003) in summarizing the research conducted by Joyce and Shovers, Wood and Thompson, and Achilles, Dickerson, Dockery-Runkel, Egelston, and Epstein cautioned district leaders from investing in single-day training. In their study of five fourth-grade teachers, Tienken and Achilles stated that “carefully planned, small-group PD implemented with an awareness of change process can be empirically tested and (a) change the instructional practices of teachers; and (b) positively impact student achievement” (p. 166).

In examining the NAEP data set in which teachers reported time spent engaging in PD activities, Tienken and Achilles (2006) noted modest evidence of the positive
educational significance of PD. "Statistically significant scale-score gains were small, generally five or fewer points, very modest given the points available (0-580) and the number of points needed to move up one category on the NAEP scale" (Tienken & Achilles, 2006, p. 14).

Other researchers stated that when implemented correctly, staff development was a key factor to student success. In his review of the Berkeley County School District Title I Project Etheridge (2001) stated:

The key to any academic intervention is the teacher...Title I resources [should be used] to support the teacher as he or she implements the approved state and district curriculum....Staff Development should result in teachers being better prepared to present the adopted curriculum, effectively supervise classroom activities, and grow professionally. (p. 341)

Dodd and Wise (2002) recommended that staff development provide time for planning and evaluation with a focus on "reviewing teaching philosophies, academic integration into the regular program and effective instructional strategies, especially with regard to the needs of academically at-risk youth" (p.2). Although there are several small-scale studies on the positive relationship of professional development on student achievement, to date there are no large-scale reputable studies that show the connection between teacher professional development and student achievement gain. The lack of SBR on the positive impact of professional development on improving student achievement was of interest to the researcher given that a portion of the Title I resources had to be allocated specifically for professional development activities.
Curriculum. Selecting the right curriculum and materials is essential for student success. Le Tendre and Chabrán (1998) commented that Title I funds should be used to buy quality curriculum materials. "Common sense and a large body of research, including TMSS, suggest that the content of curriculum and instructional materials must possess focus, depth, rigor, and meaning to improve learning in mathematics and science" (p. 308). Title I is a billion-dollar investment, so selection of quality materials, preferably research-based, is essential to run a cost-effective program.

By the mid-nineties schoolwide programs were being favored over pull-out programs, despite substantial empirical evidence supporting the effectiveness of these programs. District leaders looked for quality curriculum to enhance and enrich the abilities of all students, not just the Title I students. Researchers began to investigate the effectiveness of various curriculum programs used by Title I schools implementing schoolwide enrichment.

In a nationwide investigation, Ysseldyke, Betts, Thill and Hanagan (2004) compared the math performance of two groups of Title I students. They described the purpose of their investigation as:

We sought to ascertain the extent to which teacher use of a curriculum-based instructional management system (Accelerated Math™) as an instructional enhancement that enabled them to monitor and adapt math instruction would result in significantly greater gains in mathematics achievement for students in Title I programs than for Title I students whose teachers did not use a curricula-based instructional management system. (p. 12)
The Accelerated Math program monitored student progress toward accomplishing a set of objectives and provided teachers with information about what content to teach and how to match instruction to the skill level of the learner. An ANCOVA was used to evaluate the pre- and posttest gains of Title I students taking the STAR computer adaptive test of mathematics skills. Ysseldyke et al. (2004) found a significant difference in the Normal Curve Equivalents (NCEs) with the experimental Title I group gaining 7.9 NCEs and the control Title I group gaining 7.6 NCEs. Although these results were statistically significant, the 0.3 difference in NCEs fell far short of supporting a robust conviction that Title I students were excelling. They also conducted an ANOVA that compared the implementation factors for Title I and non-Title I students who used Accelerated Math, and found significant differences between the groups in the average percent items answered correctly and the number of math objectives mastered. Their findings were consistent with the findings of other researchers (e.g., Borman & D'Agostino, 1996) that suggested that participation in Title I programs yielded small, if any, significant gains in academic achievement and little to no gains in closing the achievement gap.

Since the onset of NCLB after-school and extended-day programs were increasingly in demand. The U.S. DOE reported a rise in the number of applications for the after-school program, 21st Century Learning Centers (De Kanter, Adair, Chung, & Storehill, 2003). These centers provided academic enrichment activities designed to help students improve academic performance in math and reading and other areas of child development. More and more after-school and extended-day programs were striving to provide academic assistance by incorporating similar to that curricula used in the daily
program, particularly those curricula that have shown results. The report from the Urban Afterschool Programs: Evaluations and Recommendations as cited in Dodd and Wise (2002) recommended that the after-school program curriculum be aligned with the regular school day curriculum and be taught by teachers who were familiar with the curriculum and had demonstrated success teaching at-risk students. The researcher would caution district leaders in aligning the after-school curriculum with the daily curriculum in absence of evidence that supports the effectiveness of the daily curriculum at improving student achievement.

*Program evaluation.* Studies done by Borman and D’Agostino (1996); Wong and Meyer (1998), Taylor and Teddie (1999), and Sunderman (2001) pointed out the importance of having implementation fidelity and effective monitoring and evaluation systems for Title I programs. Their studies indicated that in some cases failure to obtain a measurable level of success and/or the attainment of a marginal level of success of Title I programs, was in part due to the district’s lack of proper implementation, monitoring and evaluation of the programs. Feedback should be an important part of the design structure of any Title I program. Dodd and Wise (2002) reported that the U.S. Department of Education recommended the following steps to support effective program evaluation:

1. Focus the evaluation with specific questions related to program goals and strategies;
2. Identify how the program will track students’ progress;
3. Develop assessment instruments and data collection techniques;
4. Summarize data;
5. Analyze and interpret information;
6. Summarize and act on findings. (p. 25)

So how does a district leader obtain evidence of implementation fidelity of an after-school program? The same way a principal obtains evidence of implementation fidelity of daily instruction – observations and data analysis. Principals and program supervisors should conduct regular walk-throughs for the purpose of program implementation evaluation. Davidson-Taylor (2002) stated that the focus of the walk-through should be on student work and behavior. The researcher would add that student work and behavior should be viewed as a function of instructional strategies selected by the teachers. One of the goals of the Project Excel program was to meet the individual needs of the students. The selection of strategies was crucial to aligning instruction and learning. Skrenta and Fisher (2002) discussed the development of a walk-through form that would provide administrators a tool for recording the alignment of instruction and learning to a set of predetermined goals.

Despite the wealth of qualitative data obtained from a walk-through, the researcher was unable to obtain such valuable observable data due to restrictions imposed by the Institutional Review Board (IRB) at the affiliated university. Thus the researcher used past anecdotal data written by program participants and instructors as useful sources of information regarding implementation fidelity.

Opportunity to learn. Common sense dictates that in order for students to achieve they must have appropriate opportunities to learn...defined by a narrow set of instructional components” (Schwartz, 1995, p.1). Many of the components outlined in
The theoretical framework of the present study are reflected in the Opportunity to Learn (OTL) strategies. These strategies were originally proposed by the International Association for the Evaluation of Educational Achievement (IEA) as a means to "determine whether cross-national differences in students' mathematics achievement were caused by differences in students' learning experiences rather than in their ability to master the subject" (Schwartz, 1995, p. 2).

The OTL concept has evolved in the last 20 years. The federal government added additional requirements to the OTL concept to ensure that equitable education was received by disadvantaged and minority students. Schwartz (1995) stated:

The Hawkins-Stafford Education Amendments of 1988 mandated the development of OTL indicators to measure the effectiveness of Federally-funded educational programs. The resulting report by the Special Study Panel on Education Indicators (SSPEI, 1991) included a range of measurable indicators that covered both classroom experience and the overall school environment. (p. 2)

By 1992, the OTL concept was expanded to include instructional delivery standards. The National Council on Education Standards and Testing (NCEST) highly recommended the incorporation of OTL standards as a means of closing the SES achievement gap. With the Clinton Administration's Goals 2000: Educate America Act, public education leaders saw the design of OTL standards in the form of: (a) resource standards, (b) curriculum delivery standards, and (c) outcome and capacity building standards (Denver, Grant, & Jackson, 1994, p. 1). By 1999, the Federation for Children with Special Needs, Inc. identified nine OTL requirements:

(1) curriculum modified to achieve standards;
(2) instructional materials and methods;
(3) class size and structure;
4) individual assistance;
(5) supportive services;
(6) teacher training;
(7) professional development;
(8) adequate funding;
(9) teacher self-assessments and peer reviews. (Boundy, 1999, p. 1)

As the OTL variables increased in number and specificity, the demand for high quality
Title I-funded programs that ensure academic success also increased.

Summary

Chapter 2 provided a summary comparison of relevant research on Title I
programs. The research was inconclusive as to the effectiveness of Title I programs at
improving academic achievement. Several researchers reasoned that the addition of Title
I programs led to increased academic achievement of low-income students and others
found no significant difference in the performance of students participating in Title I
programs and that of those who did not participate. The research also suggested that
participation in Title I programs had little effect at closing the achievement gap.

However, obtaining an accurate assessment as to the reasons for success or failure of
Title I programs may be limited because many of the earlier programs lacked feedback
systems needed to thoroughly evaluate the programs while in existence. Much of the
data collected were based on perception rather than from empirical data. The researcher
aimed to collect empirical data on the existing Title I-funded extended-day programs offered at Riverview schools.

Chapter 3 provides an outline of the methodology the researcher used to conduct the present study. A description of the participants, length of study, the research design, instrumentation, and statistical analysis of data are included. Details of the instructional program Project Excel, a Title I-funded extended-day program, are provided as well as are discussions of data collection and data analysis procedures.
Chapter III

Methodology

Introduction

“The 50th anniversary of the Supreme Court’s school desegregation [decision, sic, Brown v. Topeka, 1954] has intensified public awareness of the persistent gap in academic achievement between [B]lack and [W]hite students” (Rothstein, 2004b, p. 1) and prompted political action and legislation that mandated that the public schools do something to eliminate the gap. School and district leaders struggled to find ways to use Title I funds effectively to support programs and initiatives to improve student achievement. The researcher’s purpose for this study was to examine the relationship between participation in Project Excel, a Title I-funded extended-day program offered at Riverview schools, and the mathematics achievement of Title I-eligible students. To investigate the nature of the influence of the program participation on student outcomes the following research questions were addressed:

1. What is the significance, if any, of preexisting differences in mathematical abilities of the participating groups in this study?

2. What is the relationship between participation in Title I programs, specifically the Project Excel program, and mathematics achievement?

3. What is the relationship between participation in Title I programs, specifically the Project Excel program and closing achievement gaps between low-income
and moderate-high-income students and non-Asian minority and White students?

This chapter describes the methodology that was used to conduct the present study. Chapter 3 contains the following subsections: (a) research design and variables, (b) participants, (c) instrumentation, (d) treatment, (e) data collection and analysis, and (f) hypotheses.

Research Design and Variables

The best research design incorporates randomization; yet, school settings often do not lend themselves to effective randomization. Therefore, a different research approach is needed to conduct meaningful research in these institutions. Often the researcher is unable use randomization techniques due to the constraints of the school organization including the size of the school, the number of instructional staff members, the physical amount of space available, and the need for parental consent. Schafer (2001) stated:

Randomization of participants to treatment conditions has long been considered a powerful method of control.... However, randomization is typically unavailable to those who work in field setting because the investigator is not able to manipulate treatment conditions at the level of the individual participant. (p. 1)

It would be unprofessional to deny any qualifying student the right to participate in the extended-day program for sake of achieving randomization. Parents who believed their children would benefit from participation and were entitled to such services might challenge the local board of education that denied them this right, and could possibly win
such a court case. However, despite the lack of randomization, meaningful research can still be conducted.

Researchers (e.g., Borg & Gall, 1989; Campbell & Stanley, 1967; Gribbons & Herman, 1997; Leedy & Ormrod, 1985; Trochim, 2000) explained that a quasi-experimental research design could be used when randomization was not possible. Borg and Gall (1989) reported that this research design was most commonly used in educational studies. The quasi-experimental design was often used for nonrandomized control group pretest-posttest investigations, in which the researcher attempted to use intact groups that were similar as the treatment and no-treatment groups (Leedy & Ormrod, 1985). The quasi-experimental design is similar to true experimental pretest-posttest designs in that both the treatment group and a control group are given a pretest; the treatment group participates in the program or particular treatment, and both groups are given a posttest. The difference is however, because randomization is not used, the quasi-experimental researcher must attempt to control for any pretreatment differences between groups. Some of these variables can be controlled statistically. Schafer (2001) summarized research conducted by Pedhazur which indicated that:

[O]ne approach used in the field is to measure extraneous variables and employ statistical control with intact groups: attempting to equate them on the outcome variable(s) using one or more pretest(s), attempting to control for other variable(s) in looking at mean differences, and attempting to control for other variable(s) in looking at differences in regressions. (p. 2)

True experimental and quasi-experimental designs have three things in common: (a) the establishment of groups, either by randomization or purposeful sampling, (b) the
exposure of the experimental group to the treatment of choice, and (c) the use of an outcome or dependent variable as a way of comparing the groups and making inferences regarding the treatment. In cases where the investigator cannot control group selection and administration of a treatment, yet still seeks a design framework similar to a quasi-experimental design, a different research design such as the nonexperimental, cross-sectional, explanatory research design as described by Johnson (2001) may be employed. Kerlinger as cited in Johnson (2001) stated:

Nonexperimental research is systematic empirical inquiry in which the scientist does not have direct control of independent variables because their manifestations have already occurred or because they are inherently not manipulable. Inferences about relations among variables are made without direct intervention, from concomitant variation of independent and dependent variables. (p. 7)

In the present study, which was cross-sectional in nature, the researcher examined past records of students who participated in the Project Excel program, an extended-day mathematics program. These records which included information regarding the length of participation in the program, class size, learning objectives mastered and instructional strategies were used along with the results from district and standardized test data to describe the relationship of program participation to mathematics achievement and closing an achievement gap.

The nonexperimental, cross-sectional, explanatory research design was selected because in the present study the researcher sought to determine the significance of gains in mathematical competencies of Title I-eligible students (Group A) following their participation in Project Excel, a Title I-funded extended-day program. Two comparison
groups were used: a Title I-eligible group who did not participate in the extended-day program (Group B) and a Title I-eligible group who did not participate in the extended-day program (Group C). Students in each of the three groups were matched and grouped to the extent possible to control for extraneous variables. Using three groups allowed the researcher to ascertain any relationships between program participation and academic improvement and indications of closing an achievement gap. Pretest program data were collected from students in all three groups. Additional test data such as TerraNova national math percentile rankings were collected from all second- and fifth-grade students. NJ ASK scaled math scores were collected from all third- and fourth-grade students.

Random selection and assignment of the participants were not possible because the assignment of participants was made by the building principal and teacher committee on the basis of several criteria, including free or reduced lunch eligibility, cutoff score on the pretest, percentile ranking on the previous year's TerraNova test and level of proficiency on the previous year's NJ ASK test.

In the present study, the researcher identified student participation in the extended-day program Project Excel as the independent variable. Participation was voluntary and the researcher had little control over the total number of hours each student actually spent in the program. Dependent variables for the investigation were the DBA mathematics posttest scores, percentile rankings on the TerraNova mathematics section and the scaled NJ ASK mathematics scores. Purposeful sampling was used to lessen the effect of factors occurring during the school day that may have impacted the outcome of the study such as variable class size, presence of classroom aides, differences in math
Teachers and participation in additional schoolwide programs such as basic skills math pull-out instruction.

The nonexperimental, cross-sectional, explanatory research design allowed the researcher to evaluate the influence of student participation in the extended-day program on student achievement; and this design also served as an evaluative tool to allow "... managers of quality programs to judge the efficacy of their efforts on the basis of established, accepted goals for the program and improve their effectiveness in promoting the resilience of at-risk students" (Huang, 2001, p. 45). Program evaluation was a necessary component of Title I program design, implementation fidelity and improvement. Kemper and Achilles (1979) suggested use of an evaluation outline that covers: (a) program initiation including a needs assessment and program planning, (b) program operations including implementation and improvement, and (c) program outcomes including analysis of effect, product/comparative evaluation and cost effectiveness (p. 20).

Because this investigation was post hoc in nature, the researcher focused more on program outcomes rather than on program initiation and operations. Attendance records, student portfolio, DBA data, and standardized test data were analyzed to conduct what Kemper and Achilles (1979) described as "effect evaluation - relat[ing] actual program operations and procedures to either intended or unanticipated outcomes, or both" (p. 23). Although effect evaluation suggests determinations of cause and effect, the researcher did not attempt to establish causation; rather, the researcher sought to evaluate and understand the relationship between student participation in extended-day programs and student achievement.
Participants

The participants represented students enrolled in Riverview Schools, a pseudonym, entering Grades 2 through 5 during the 2004-05 school year. Riverview School District is a diverse community with a student population of 49% White, 20% Asian, 16% African American, and 15% Hispanic. The district, located in the heart of Central Jersey, has four schools: a primary school (PK – 2), an elementary school (3 – 6), a middle school (7 – 8), and a high school (9 – 12). Located near a major university and two teaching hospitals, Riverview is home to many visiting professors and doctors which in part explain the high mobility rate of 22%. Despite the vast amount of wealth in the community, nearly 23% of the 1500 students attending Riverview schools qualify for free or reduced lunch.

District records for free and reduced lunch were used to identify students from low-income families as the target population. Previous year's standardized test scores, district mathematics pretests, benchmark assessments, and teacher recommendations were used to identify students who qualified for additional mathematics instruction. Invitations to participate in the program were mailed to students' homes and teachers contacted the parents who did not return the written invitation. Parents voluntarily elected to allow their children to participate in the program. Students participating in the program were categorized as Group A (n = 111). Students who were invited to participate and elected not to participate were categorized as Group B (n = 92). Students in both groups were eligible to receive Title I services and were selected from primarily low-income families. A third group, Group C (n = 203) was selected from students who were not eligible for Title I services, from primarily moderate/high-income families, and
did not qualify for participation in the Project Excel extended-day program. Table 1 summarizes the sample sizes for all students participating in the study. Matched grouping was used to the extent possible to control for variables such as math teacher, basic skills intervention (BSI), variations in structure of the day, time of day and length of math period, and gender. Students who entered the district midyear and did not take the pretest were excluded from analyses regardless of participation in the Project Excel program. Likewise, students who transferred and did not take the spring standardized tests were excluded.

Table 1

<table>
<thead>
<tr>
<th>Grade</th>
<th>Participating</th>
<th>Eligible</th>
<th>Ineligible</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group A (n)</td>
<td>Group B (n)</td>
<td>Group C (n)</td>
<td>(N)</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>23</td>
<td>56</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>22</td>
<td>22</td>
<td>52</td>
<td>96</td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td>28</td>
<td>39</td>
<td>102</td>
</tr>
<tr>
<td>5</td>
<td>33</td>
<td>19</td>
<td>56</td>
<td>108</td>
</tr>
<tr>
<td>Total</td>
<td>111</td>
<td>92</td>
<td>283</td>
<td>406</td>
</tr>
</tbody>
</table>

One constraint of using a nonexperimental, cross-sectional, explanatory research design was that the researcher had no control over the selection process used to identify group populations. Although matched grouping of data were used to account for similarities in gender, daily program, math teacher, and curriculum, participation in special education programs and G&T programs as well as ethnicity were three variables
that could not be controlled for. Table 2 provides a summary of the demographics of each of the three groups.

Of the total population of students (n = 466), 50% were ineligible to participate in Title I programs, 22.7% were eligible but elected not to participate, and 27.3% were eligible and elected to participate. The population distribution of special-education students was skewed in that only 11.6% were classified as ineligible, 32.6% were classified as eligible but did not participate, and 55.8% participated. The opposite pattern existed with the G&T students in that 95% were classified as ineligible, 5% were classified as eligible but opted not to participate, and 0% participated. Of the eight second-grade special-education students, six participated in Project Excel, one student elected not to participate, and one student was not eligible. Of the 17 G&T second-grade students, only two were eligible and both elected not to participate. Both the second-grade special-education and G&T students received additional mathematics instruction during the day as a part of a pull-out program.

Of the nine third-grade special-education students, two participated in Project Excel program, three opted not to participate, and the remaining four students were ineligible. All 14 G&T third-grade students were ineligible. The third-grade special education students received additional mathematics instruction during the day from both a regular education teacher and a special education teacher as a part of a pull-out program. The third-grade G&T students received 30 minutes of additional enrichment mathematics instruction twice a week during the day.

All 18 fourth-grade special-education students were Title I-eligible but only 9 elected to participate in the Project Excel program. Of the 16 G&T fourth-grade students,
Table 2

Demographics of Student Population

<table>
<thead>
<tr>
<th>Variable</th>
<th>Participating</th>
<th>Eligible</th>
<th>Ineligible</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group A (n = 111)</td>
<td>Group B (n = 92)</td>
<td>Group C (n = 203)</td>
<td>(N = 406)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>48</td>
<td>50</td>
<td>105</td>
<td>203</td>
</tr>
<tr>
<td>Female</td>
<td>63</td>
<td>42</td>
<td>98</td>
<td>205</td>
</tr>
<tr>
<td>Classification</td>
<td>87</td>
<td>78</td>
<td>122</td>
<td>283</td>
</tr>
<tr>
<td>General Ed.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Special Ed.</td>
<td>24</td>
<td>14</td>
<td>5</td>
<td>43</td>
</tr>
<tr>
<td>G &amp; T</td>
<td>0</td>
<td>4</td>
<td>76</td>
<td>80</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>38</td>
<td>38</td>
<td>132</td>
<td>208</td>
</tr>
<tr>
<td>Black</td>
<td>34</td>
<td>19</td>
<td>10</td>
<td>63</td>
</tr>
<tr>
<td>Hispanic</td>
<td>33</td>
<td>24</td>
<td>8</td>
<td>65</td>
</tr>
<tr>
<td>Asian</td>
<td>6</td>
<td>11</td>
<td>53</td>
<td>70</td>
</tr>
</tbody>
</table>

only one was eligible and elected not to participate. The special-education students received additional mathematics instruction during the day from both a regular education teacher and a special education teacher as a part of a pull-out program. The fourth-grade G&T students received accelerated instruction using the fifth-grade math curriculum in lieu of attending the regular fourth-grade math classes with their peers. During the hour
of mathematics instruction the G&T fourth graders proceeded to another room to be instructed by the G&T math teacher.

All eight fifth-grade special-education students were Title I-eligible, seven participated in Project Excel, and one student elected not to participate. Of the 33 G&T fifth-grade students, only one was eligible and elected not to participate. The special-education students received additional mathematics instruction during the day as a part of a pull-out program. The fifth-grade G&T students received enrichment instruction using the sixth-grade math curriculum in lieu attending the regular fifth-grade math classes with their peers. During the hour of mathematics instruction the G&T fifth graders traveled to another room to be instructed by the G&T math teacher.

Inequity in group assignments and differences in daily mathematics instructional focus and time resulted in the special-education and G&T students being excluded from the present study. Table 3 summarizes the modified sample information for this study. Although the ethnic distributions were also skewed with a larger percentage of Asian and White students classified as ineligible and larger percentage of Black and Hispanic students classified as participating, when the special education and G&T students were excluded, all four ethnic subgroups received the same amount of daily mathematics instruction during the day. This allowed the researcher to compare the groups for evidence in closing the achievement gap.

**Instrumentation**

The researcher used several tools to measure student performance including district grade-level benchmark assessment mathematics pretests, posttests in the form of
Table 3

**Final Study Sample -- Special Education and G&T Students Excluded**

<table>
<thead>
<tr>
<th>Grade</th>
<th>Participating</th>
<th>Eligible</th>
<th>Ineligible</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group A (n)</td>
<td>Group B (n)</td>
<td>Group C (n)</td>
<td>(n)</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>20</td>
<td>40</td>
<td>75</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>19</td>
<td>34</td>
<td>73</td>
</tr>
<tr>
<td>4</td>
<td>26</td>
<td>19</td>
<td>24</td>
<td>69</td>
</tr>
<tr>
<td>5</td>
<td>26</td>
<td>17</td>
<td>24</td>
<td>67</td>
</tr>
<tr>
<td>Total</td>
<td>87</td>
<td>75</td>
<td>122</td>
<td>284</td>
</tr>
</tbody>
</table>

Second- and fifth-grade TerraNova mathematics tests, and Grades three and four NJ ASK tests, and program participation student performance sheets.

All students in the district were assessed using District Benchmark Assessments or DBAs. The DBAs, given in mathematics, were criterion-referenced tests created by Everyday Learning Corporation, modified by a mathematics committee and aligned with the New Jersey Core Curriculum Content Standards (NJCCCS) for each grade level. The battery of assessments consists of four components: a pre-assessment given in September, a midyear assessment given in January, a post assessment given in June, and nine monthly benchmark assessments. Each item on the individual assessments was linked to a learning goal for a particular unit taught in the Everyday Mathematics program. Everyday Mathematics was the curriculum used in all first through fifth-grade classes with the exception of the G&T fifth-grade class which uses the Connected Math curriculum. The learning goals in Everyday Mathematics are aligned to the NJCCCS.
Appendix A includes a summary of the learning goals addressed at each grade level, and Appendix B outlines the NJCCCS for mathematics. The benchmark assessment test items included short-constructed responses, fill-in-the-blank, multiple choice, matching, and open-ended questions.

Identical test items were included on all grade-level pre- and posttests. Classroom teachers administered the pretest, midyear, and post assessments to their students during their hour of math instruction, received training on the use of benchmark assessment rubrics, and indicated student growth as per district guidelines. Teachers were given a rubric to score the pretests and scores were reported as the percentage items answered correctly. Pretests with fewer than 60% items answered correctly were one criterion used for placement in the Project Excel program. Pre- and posttest data were recoded by a confidential secretary using the students' identification numbers and given to the researcher for comparative analysis. However, in examining the coded data, the researcher learned that not all teachers administered the post assessment in June as directed; therefore, a modified examination of posttest scores is included in this study.

Data from the third- and fourth-grade New Jersey ASK mathematics tests, which are criterion-referenced tests aligned to the NJCCCS for mathematics, were also recoded by the confidential secretary and given to the researcher for comparative analysis. The NJ ASK has two major types of questions: multiple choice and open-ended. "Multiple choice questions add much to the reliability and consistency of the test because many good questions that focus on a broad range of skills can be answered in a short span of time" (New Jersey Professional Educators Port PowerPoint slide 11, 2005). The open-ended question "allows the children to express what they know about each question in
their own words. Students may also present their response using diagrams, graphics, and/or pictures” (New Jersey PEP PowerPoint slide 12, 2005). Students received a scaled score out of a possible 306 in which 250 and higher was considered advanced proficient, 208 to 249 was proficient, and less than 200 was partially proficient. The researcher gathered data on students’ scaled scores as well as their raw scores on the five mathematics clusters: (a) number sense and operations, (b) geometry and spatial sense, (c) algebra and patterns, (d) data analysis and discrete math, and (e) problem solving.

NJ ASK performance was a second criterion used for placement in the Project Excel program. Students scoring in the partially proficient and low proficient range were recommended to participate in the program. Tienken (2006) cautioned district leaders from relying on the NJ ASK as a diagnostic tool (personal communication, January 23, 2006). Referring to the data presented in the 2004 NJ ASK Technical Report, Tienken stated, “I would caution people from basing program decisions on the results of these assessments [NJ ASK and GEPA], ... they have high standard error of measure and low reliability in the cluster scores. They are very blunt instruments” (Tienken, personal communication, January 23, 2006). Table 4 shows the reliability estimates and standard errors of measurement (SEM) for the NJ ASK. Although in the present study the researcher compares the NJ ASK mathematics performance of Project Excel participants to that of various groups, the findings must be interpreted with caution due to reliability and SEM concerns with the instrument.

Validity issues were also of concern. An excerpt from the NJ DOE 2004b technical report read “…an external committee is assisting the New Jersey Department of Education by reviewing the assessments to determine how well they measure the
knowledge and skills stated in the standards" (p. 59). District leaders had no assurances that what was tested matched what was taught based on the standards. This presents another reason why district leaders should proceed cautiously when using the NJ ASK as a diagnostic tool for remediation. A detailed description of the NJ ASK tests can be found in Appendix C.

The TerraNova Test was given in Grades 2 and 5 and the data were rescoded by a confidential secretary and provided to the researcher. The TerraNova, similar to the California Achievement Test (CAT), Metropolitan Achievement Test (MAT) and the Comprehensive Test of Basic Skills (CTBS) is a norm-referenced test in which scores can be reported as percentiles. Students' overall national mathematics percentile rankings as well as their individual performance (scale of 0 - 100) on each of the individual mathematics objective performance indexes (OPI) were examined and compared. There were eight OPIs on the second-grade TerraNova including: (a) number and numerical relationships, (b) computation and estimation, (c) geometry and spatial sense, (d) data analysis, statistics, and probability, (e) problem solving and reasoning, (f) communication, (g) addition of whole numbers, and (h) subtraction of whole numbers.

The fifth-grade TerraNova also included eight OPIs, the first six of which were identical to the first six indicators used on the second-grade assessment. Addition and subtraction of whole numbers were replaced with measurement and operation concepts on the fifth-grade test. The OPI is based on a minimum of four test items and provides an estimate of the number of items a student could theoretically solve if there had been 100 such items for that objective. Students scoring an OPI of 75 to 100 were at the mastery
### Table 4

Reliability Estimates and Standard Errors of Measurement (SEM) for Mathematics

**Clusters – 2004 NJ ASK Grades 3 and 4**

<table>
<thead>
<tr>
<th>Grade</th>
<th>Cluster</th>
<th>Raw Points</th>
<th>Reliability Estimates</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 4</td>
<td>Number Sense &amp; Numerical Operations</td>
<td>13</td>
<td>.73</td>
<td>1.47</td>
</tr>
<tr>
<td></td>
<td>Geometry &amp; Measurement</td>
<td>10</td>
<td>.59</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>Patterns &amp; Algebra</td>
<td>10</td>
<td>.69</td>
<td>1.71</td>
</tr>
<tr>
<td></td>
<td>Data Analysis, Probability &amp; Discrete Math</td>
<td>10</td>
<td>.61</td>
<td>1.62</td>
</tr>
</tbody>
</table>

Grade 3

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Raw Points</th>
<th>Reliability Estimates</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Sense &amp; Numerical Operations</td>
<td>9</td>
<td>.71</td>
<td>1.10</td>
</tr>
<tr>
<td>Geometry &amp; Measurement</td>
<td>5</td>
<td>.41</td>
<td>.86</td>
</tr>
<tr>
<td>Patterns &amp; Algebra</td>
<td>5</td>
<td>.59</td>
<td>.93</td>
</tr>
<tr>
<td>Data Analysis, Probability &amp; Discrete Math</td>
<td>5</td>
<td>.59</td>
<td>.90</td>
</tr>
</tbody>
</table>

*Note: Source: NJ DOE, 2004b*

At the grade level, students scoring an OPI of 50 to 74 were at the partial mastery level, and students scoring an OPI of less than 50 were at the non-mastery level.

The national percentile ranking obtained from the mathematics composite score was another criterion used to identify Title I-eligible students. The TerraNova had a higher reliability and a lower SEM relative to the number of questions asked than did the
NJ ASK tests. The increased reliability perhaps made the TerraNova a more useful diagnostic tool than the NJ ASK; however, generic labeling of categories as mathematics, mathematics computation, and mathematics composite made it difficult for district leaders to identify student strengths and weaknesses in the four main mathematics clusters (Table 5). District leaders should use care when using the TerraNova as a diagnostic tool for remediation.

Qualitative data needed to provide information on implementation fidelity were obtained from individual performance sheets of the students participating in the program. Portfolios contained student attendance records, learning objectives taught, instructional methodologies used, and progress made towards achieving the learning goals. Information collected from teacher time sheets was used to verify teacher attendance. All information was recorded by the confidential secretary.

The Program

Project Excel was an after-school program that provided mathematics remediation to students identified as skills-fragile. The program, which was fully funded by Title 1 monies, employed 15 teachers in two buildings. Each teacher received 2 hours of initial training which focused on strategies for working with at-risk students, forms and procedures for recording student progress, roles and responsibilities, parent and teacher communication, and goals and expectations for the program. At the conclusion of the program, teachers attended another 2-hour workshop to evaluate the program informally, discuss successes, and make plans for implementing changes to the program the following year. Over 125 students in Grades 1 through 5 participated.
Table 5

Reliability Estimates and Standard Errors of Measurement (SEM) for TerraNova Spring Form A Multiple Assessments

<table>
<thead>
<tr>
<th>Grade 2</th>
<th>Raw Points</th>
<th>Reliability Estimates</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics</td>
<td>44</td>
<td>.85</td>
<td>3.06</td>
</tr>
<tr>
<td>Mathematics Computation</td>
<td>20</td>
<td>.84</td>
<td>1.72</td>
</tr>
<tr>
<td>Mathematics Composite</td>
<td>64</td>
<td>.91</td>
<td>3.52</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade 5</th>
<th>Raw Points</th>
<th>Reliability Estimates</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics</td>
<td>43</td>
<td>.92</td>
<td>3.82</td>
</tr>
<tr>
<td>Mathematics Computation</td>
<td>20</td>
<td>.83</td>
<td>1.78</td>
</tr>
<tr>
<td>Mathematics Composite</td>
<td>63</td>
<td>.93</td>
<td>4.29</td>
</tr>
</tbody>
</table>

*Note. Source: CTB Mc-Graw-Hill, 2001*

regularly in the program; first-grade participants were excluded from this study due to lack of corresponding standardized testing data. Children attending the second-grade extended-day program met 60 minutes after school one day each week. The program was designed to provide these students 45 minutes of instruction in which the teachers used direct instruction, guided practice and hands-on activities, and 15 minutes of performance-based assessment each week.
Children attending the third- through fifth-grade program met 75 minutes after school one day each week. The program was designed to provide children with an average of 20 minutes of computer-assisted instruction, 40 minutes of hands-on and direct mathematics instruction, and 15 minutes of performance-based assessment each week. Project Excel teachers worked with three to five students per day, and teachers elected to teach one to four sessions per week. Students were grouped according to grade level and regular classroom teacher, and were assigned to one Project Excel teacher for the duration of the program. The program ran from October through May.

A flexible individualized learning portfolio was developed for each student based on their needs. The Project Excel teacher worked closely with the regular classroom math teacher to develop the portfolio. Information from the standardized tests, pretests, and monthly benchmark assessments was used to develop the portfolio. Project Excel teachers with the assistance of the regular math teacher, identified areas of weakness and designed programs to reinforce specific learning goals.

During the computer-assisted instruction time children worked independently on activities through web-based programs such as www.aas.math.com and www.mathgoodies.com and CD-ROM programs such as Math for the Real World. During the hands-on and direct instruction time students completed selected activities from the following programs: Measuring Up to New Jersey Core Content Standards and Everyday Mathematics. Exemplars, open-ended questions, selected games, and hands-on activities were used during the direct instruction and performance-based assessment times.
Project Excel teachers held monthly meetings with regular classroom teachers to discuss the progress of the students. Project Excel teachers communicated with parents monthly by providing them with progress reports and scheduling conferences whenever necessary. Parents were regularly contacted via telephone if their child was absent from the program.

Data Collection and Analysis

With the approval of the superintendent, archival data were gathered from the school district's computerized database. A confidential secretary recoded all student data to preserve student anonymity prior to giving data to the researcher. Data included standardized test scores, demographic data, information regarding free or reduced lunch status, participation status in programs such as special education, G&T, and Project Excel, math teacher assignments, attendance records, mathematics benchmark assessment scores, and student performance sheets. The pretest given in September was a criterion-referenced test, and students who scored below 60% of the items answered correctly met the initial selection criteria for participation in Project Excel. Pretest and posttest items included numeracy and operations; geometry and spatial sense; patterns and algebra; and data analysis, probability, and discrete math. It was anticipated that data from the posttest given in June would be compared to the pretest scores. However, not all teachers administered the posttest as directed; thus, only pre-and posttest gain scores for second- and third-grade students were included in this study.

The driving force behind Title I was the improvement of academic achievement of low-income students and closing of the achievement gap between these students and
their economically stable counterparts. One cannot comment as to whether these gaps were closing without first establishing the existence of the gaps in absence of the treatment. To do so, the researcher compared the mean pretest scores of Project Excel participants \((n = 87)\) with those of Title I-eligible non-participants \((n = 75)\) and all non-participants \((n = 197)\). All students completed the pretest, therefore the researcher did not have to partition the sample based on the spring assessment taken. To examine the differences in light of the minority achievement gap, the researcher compared the mean pretest scores of non-Asian minority participants \((n = 53)\) with that of White non-participants \((n = 111)\) to determine the existence and significance of a minority achievement gap prior to participation in the Project Excel program. In a similar investigation, the researcher compared the mean pretest scores of lower SES participants \((n = 56)\) with those of moderate/higher SES non-participants \((n = 146)\) to determine the existence and significance of an SES achievement gap prior to participation in the Project Excel program. These two investigations assisted the researcher in addressing the first research question: What is the significance, if any, of preexisting differences in mathematical abilities of the participating groups in this study? A graphic representation of the first investigation and subsidiary questions can be seen in Figure 2.

One of the first questions posed by district leaders after the Project Excel program began was how well did the students perform? This question required the researcher to establish a baseline for comparing the math achievement of the participating students to the rest of the student body. The researcher used both quantitative and qualitative data collection methods. Quantitative methods included comparison of the mean performance of the Project Excel participants (Group A) on the TerraNova math test, NJ ASK math
Figure 2. Graphical representation of the first research question. What is the significance, if any, of preexisting differences in mathematical abilities of the participating groups in this study?
test, and DBA math posttest with the mean performance of two groups: (a) the Title I-eligible non-participants (Group B) and (b) all of the students who did not participate in the study (Groups B and C combined). Earlier studies on the effectiveness of Title I programs were often criticized because of lack of a control group; hence, for the purpose of investigating the first research question the researcher sought to include two comparison groups: all of the non-participating students (Groups B and C) and a group of students with similar mathematical backgrounds, Title I-eligibility, and SES status (Group B).

The use of different end-of-the-year assessments at each grade level made it difficult for the researcher to analyze the entire population of Title I participants and Title I non-participants statistically. Group A ($n = 87$) and Group B ($n = 75$) respectively; therefore, the groups were subdivided based on the spring assessments given at each grade level. Second- and third-grade students were the only students to complete both the mathematics pre- and posttests, thus the researcher used a modified sample from Group A ($n = 35$), Group B ($n = 39$), and Group C ($n = 74$) to do a pre- and posttest analysis. Second- and fifth-grade students took the TerraNova during the spring, hence the researcher modified the sample size accordingly, Group A ($n = 41$), Group B ($n = 37$), and Group C ($n = 64$) to do an analysis of TerraNova mathematics percentile rankings. The NJ ASK was given to third- and fourth-grade students; therefore, the researcher modified the sample size to do an analysis of NJ ASK mathematics performance that included 46 students in Group A, 38 students in Group B, and 58 students in Group C.

The researcher was concerned that subdividing the sample would yield small sample sizes that might increase the probability of Type I and Type II errors. Hinkin,
Wiersma and Jurs (2003) suggested using a power value of 0.80 to decrease the probability of encountering the more serious Type I errors. Using statistics tables and a power value of 0.80, a significance level of 0.05 and a median effect size of 0.75, the appropriate minimum sample size was determined to be 29 in each group (Hinkle, Wiersma & Jurs, 2003, p. 654). Having met the minimum required sample size for Groups A (n = 46), B (n = 38) and C (n = 58), the researcher minimized the probability of encountering Type I and Type II errors.

Collection and analysis of these quantitative data enabled the researcher to address the second research question: What is the relationship between participation in Title I programs, specifically the Project Excel program, and mathematics achievement? To investigate the relationship between Project Excel participation and mathematics achievement the researcher proposed three subsidiary questions: (1) How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scores of Project Excel participants compare with those of all non-participants? (2) How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scores of Project Excel participants compare with the district, state, and national norms? and (3) How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scores of Project Excel participants compare with those of the Title I-eligible non-participants? A graphic representation of the second investigation and subsidiary questions can be seen in Figure 3.

To address this first subsidiary question the researcher used independent samples t-tests to compare the DBA mathematics posttest scores, TerraNova mathematics national
Figure 3. Graphical representation of the second research question. What is the relationship between participation in Title I programs, specifically the Project Excel program, and mathematics achievement?
percentile rankings, and NJ ASK mathematics scaled scores of Project Excel participants (Group A) to those of all non-participants (Groups B and C). Where pre- and posttesting were involved Parker and Gerber (2000) used correlated t tests as their statistical tool of evaluation. They examined the pre- and posttest scores of 11 students who participated in a 5-week summer science intervention program. Although the statistical results of the correlational matched-pair t test yielded a statistically significant correlation, Parker and Gerber’s investigation incorporated a small sample size and lacked a comparable control group which in turn weakened their assertion that participation in the intervention program had an effect on improving academic achievement. Such gains could have been attributed to small class size.

In a separate study Edwards et al. (2001) used matched-pair t tests to analyze the influence of participation in a seventh- and eighth-grade summer math program. The t test results suggested there was significant difference in the posttest scores when compared to the pretest scores, thereby suggesting academic improvement. Although Edwards et al. used an average sample size of 36 in each group, they neglected to incorporate a control group thus subjecting their conclusions to scrutiny.

In their evaluation of a career intervention on at-risk middle school students’ academic achievement, Legan and Hoare (2004) collected and analyzed pre-and posttest data using a t test which incorporated a control group. At-risk students (n = 27) participated in the career intervention and the result of their academic achievement on a posttest was compared to that of 30 at-risk students who did not participate in the career intervention program. Their findings revealed that there were no statistical differences between the control and experimental groups. Despite the implementation challenges
faced by Parker and Gerber (2000) and Edwards et al. (2001), and the lack of statistical significance found by Legum and Hoare (2004), t tests are effective tools to use when comparing two independent groups provided that “both underlying populations are normally distributed with equal variances” (Witte & Witte, 2004, p. 336).

To try to avoid some of the problems encountered by previous investigators, the researcher sought to incorporate a control (or at least a strong comparison) group to compare the various test scores of groups of students who had similar mathematical backgrounds, were eligible to receive Title I services, and had similar SES. While use of a control group would provide the researcher with valuable data regarding the relative performance of Title I-eligible students, it did not address the question of how did these students’ math performance compare to that of the rest of the student population who were taking the same tests. A comparison of the mean performance of participants (Group A) compared to the mean performance of the rest of the student population was needed.

The present study had as one of its underlying premises the assumption that students selected to participate in the Project Excel program were chosen because they had not demonstrated mathematical proficiency on the previous assessments; therefore, differences in their mathematical ability level as compared to their ineligible counterparts already existed. In several cases where pre- and posttest designs were used, researchers used Analysis of Covariance (ANCOVA) to make a statistical adjustment for these pretest differences between groups. In a pretest-posttest quasi-experimental research Birch (2002) used ANCOVA to compare the mathematics achievement of students who participated in computer-assisted instruction to those students who did not participate.
The use of ANCOVA allowed him to control partially for differences between the groups that existed before he introduced computer-assisted instruction. Borg and Gall (1989) stated that ANCOVA was another statistical procedure that could be used to control for differences in prior knowledge between the two groups. ANCOVA techniques were also used by Interactive, Inc. and Metis Associates, Inc. (2002) in their multiyear evaluation of the Cleveland Municipal School District's implementation of the Lightspan Achieve Now program. They compared the performance on the spring 2002 SAT-9 reading and math tests of three groups of students while controlling for pretreatment differences in the spring 2001 OP1/OOPT scores.

Dinger, Ogleece, and Johnson (2000) conducted a quasi-experimental investigation to impact of participation in a health education curriculum. They used a nonequivalent pretest-posttest design and "because significant differences existed in pretest scores...researchers decided to use the pretest score as a covariate...ANCOVA was calculated to test for posttest differences among the four groups, while reducing the effects of initial group pretest score differences" (p. 1). Researchers Turpin and Cage (2004) also used a quasi-experimental nonequivalent control group design to analyze the influence of participation in an activity-based science curriculum on student achievement, science process skills, and science attitudes. An ANCOVA was conducted on the seventh-grade ITBS scores using the sixth-grade ITBS scores as a covariate.

Although ANCOVA is a popular statistical tool to use to control for pretreatment differences between groups, particularly pretest scores, one must be cautious not to violate several major underlying assumptions of its use. The first assumes the relationship between the dependent variable and the covariate is linear. In the case of the
present investigation, a linear relationship existed when pretest scores were coded as the covariate and TerraNova percentile rankings were coded as the dependent variable. Similar cases of linearity existed when NJ ASK math scores and the DBA math posttest scores were coded as the dependent variables. So why not use ANCOVA as the statistical tool of choice? The answer lies within the second underlying assumption.

The second assumption, often called homogeneity of regression assumed the regression lines for the individual groups are both linear and parallel (Hinkle, Wiersma & Jurs, 2003). An analysis of the data collected in the present study yielded non-parallel regression lines, thus prompting the researcher to select a more appropriate statistical tool. A third consideration was the question of actual sample size. With relatively small samples, this researcher chose to use methods of analysis other than ANCOVA.

The second subsidiary question asked: How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scores of Project Excel participants compare with the district, state, and national norms? Wong and Meyer (1998), Borman and D'Agostino (1996), and Clayton (1991) suggested the use of normal curve equivalent (NCE) gain scores as means of comparing national percentile rankings. Etheridge (2001) compared the change in percentage of students performing below the 25th percentile and students performing above average, above the 50th percentile. However, the organization of Riverview District was such that the TerraNova test was first given in the spring of grade two and then not administered again until the spring of grade five. The infrequency of TerraNova administration made it difficult to examine gains in NCE percentile rankings from year to year.
When using norm-referenced data, other researchers have used ANOVA to compare the test score means of two or more groups. McGlaughlin, Knoepf, and Holliday (2005) used ANOVA to analyze the outcome of the Woodcock-Johnson III, a norm-referenced achievement test, as it related to several factors that influenced the mathematical achievement of college students. However, ANOVA assumes the samples must be random and no significant pretreatment differences exist (Hinkle, Wiersma & Jurs, 2003). Thus, to address the second subsidiary question, the researcher used a one-sample t test to compare the TerraNova mean percentile rankings of the second- and fifth-grade Project Excel participants to the national norm for the TerraNova mathematics test. A similar one-sample t test was used to compare the NJ ASK mathematics scores of third- and fourth-grade Project Excel participants to the state average on the NJ ASK mathematics test. There were no state or national norms for the DBA mathematics posttest. Therefore the researcher used a one-sample t test to compare the posttest average of the participants to the district posttest average. Comparing the performance of Project Excel participants to district, state, and national norms allowed the researcher to gain insights as to how these students performed compared to a larger population.

The third subsidiary question asked: How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scores of Project Excel participants compare with those of the Title I-eligible non-participants? To address this question, the researcher used independent samples t tests to compare the mathematics performance of Project Excel participants (Group A) on all three assessments to that of Title I-eligible non-participants (Group B). Table 6 provides a summary of the investigations designed to address the second research question: What
is the relationship between participation in Title I programs, specifically the Project Excel program, and mathematics achievement?

Table 6

Summary of Subsidiary Questions and Statistical Analysis Associated with Research Question Two

Subsidiary Question 2-A: How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scores of Project Excel participants compare with those of all non-participants?

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Groups Compared</th>
<th>Data Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBA Posttest</td>
<td>Participants (Group A, n = 35)</td>
<td>Independent</td>
</tr>
<tr>
<td>2nd &amp; 3rd Grade</td>
<td>Non-participants (Groups B &amp; C, n = 111)</td>
<td>samples t test</td>
</tr>
<tr>
<td>TerraNova</td>
<td>Participants (Group A, n = 41)</td>
<td>Independent</td>
</tr>
<tr>
<td>2nd &amp; 5th Grade</td>
<td>Non-participants (Groups B &amp; C, n = 101)</td>
<td>samples t test</td>
</tr>
<tr>
<td>NJ ASK</td>
<td>Participants (Group A, n = 46)</td>
<td>independent</td>
</tr>
<tr>
<td>3rd &amp; 4th Grade</td>
<td>Non-participants (Groups B &amp; C, n = 96)</td>
<td>samples t test</td>
</tr>
</tbody>
</table>
Table 6 (continued)

Subsidiary Question 2-B: How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scores of Project Excel participants compare with the district, state, and national norms?

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Groups Compared</th>
<th>Data Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBA Posttest</td>
<td>Participants (Group A, n = 35)</td>
<td>One-sample t test</td>
</tr>
<tr>
<td>2nd &amp; 3rd Grade</td>
<td>District DBA Posttest Average</td>
<td></td>
</tr>
<tr>
<td>TerraNova</td>
<td>Participants (Group A, n = 41)</td>
<td>One-sample t test</td>
</tr>
<tr>
<td>2nd &amp; 5th Grade</td>
<td>TerraNova National Percentile Norm</td>
<td></td>
</tr>
<tr>
<td>NJ ASK</td>
<td>Participants (Group A, n = 46)</td>
<td>One-sample t test</td>
</tr>
<tr>
<td>3rd &amp; 4th Grade</td>
<td>NJ ASK State Average</td>
<td></td>
</tr>
</tbody>
</table>
Table 6 (continued)

Subsidiary Question 2-C: How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scores of Project Excel participants compare with those of the Title I-eligible non-participants?

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Groups Compared</th>
<th>Data Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBA Posttest</td>
<td>Participants (Group A, n = 35)</td>
<td>Independent</td>
</tr>
<tr>
<td>2nd &amp; 3rd Grade</td>
<td>Title I-eligible non-participants (Group B, n = 39)</td>
<td>samples t test</td>
</tr>
<tr>
<td>TerraNova</td>
<td>Participants (Group A, n = 41)</td>
<td>Independent</td>
</tr>
<tr>
<td>2nd &amp; 5th Grade</td>
<td>Title I-eligible non-participants (Group B, n = 37)</td>
<td>samples t test</td>
</tr>
<tr>
<td>NJ ASK</td>
<td>Participants (Group A, n = 46)</td>
<td>Independent</td>
</tr>
<tr>
<td>3rd &amp; 4th Grade</td>
<td>Title I-eligible non-participants (Group B, n = 38)</td>
<td>samples t test</td>
</tr>
</tbody>
</table>

To answer the achievement gap question effectively, the researcher had to establish comparison groups along two lines: racial and socioeconomic status (SES). Figure 4 provides a graphical representation of the third investigation. Data from the last 30 years of National Assessment of Educational Progress (NAEP) suggested that the racial achievement gap in mathematics between Whites and non-Asian minorities was
Figure 4. Graphical representation of the third research question. What is the relationship between participation in Title I programs, specifically the Project Excel program, and closing the achievement gaps between low-income and moderate/high-income students and non-Asian minority and White students?
steadily closing until the late 1980s, yet has been widening over the last decade (North Central Regional Educational Library, 2005). Disaggregated data from recent New Jersey state assessments reflected the existence of two predominant achievement gaps, a minority achievement gap, and an SES achievement gap. In the present investigation, data were disaggregated and recoded to allow the researcher to examine the minority achievement gap by computing the math achievement of non-Asian minority participants, all of whom were Black or Hispanic to that of White non-participants. This comparison required the researcher to modify the sample size by cross-referencing the groups based on ethnicity and spring assessment taken. Cross-referencing allowed the researcher to compare the DBA posttest performance of non-Asian minority second- and third-grade Project Excel participants (n = 24) with that of second- and third-grade White non-participants (n = 65). In a similar investigation using the TerraNova math percentile rankings, the researcher modified the sample size to include 28 non-Asian minority Project Excel participants and 60 White non-participants in Grades 2 and 5. The third- and fourth-grade NJ ASK math performance of 25 non-Asian minority Project Excel participants was compared to that of 51 White non-participants. In assessing the impact of participation in Project Excel at reducing the minority achievement gap, the following subsidiary questions were addressed:

3-A How do the DBA mathematics posttest scores, TerraNova mathematics national percentile ranking, and NJ ASK mathematics scaled scores of non-Asian minority Project Excel participants compare with those of White non-participants?
3-B How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scaled scores of non-Asian minority Project Excel participants compare with the district, state, and national norms?

Data regarding participation in free and reduced lunch programs were used to identify lower SES students. This data allowed the researcher to compare the mathematics achievement means of lower SES participants with those of moderate/higher SES non-participants. The researcher partitioned the sample sizes by cross-referencing SES with spring assessment taken. The researcher proposed the following subsidiary questions to address the SES achievement gap:

3-C How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scaled scores of lower SES participants compare with those of moderate/higher SES non-participants?

3-D How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scaled scores of lower SES participants compare with the district, state, and national norms?

A summary of the achievement gap investigations is presented in Table 7.

Independent t tests were used to analyze the impact of participation at reducing the racial and SES achievement gaps. The researcher was concerned about the effect of small sample sizes due to the small population of non-Asian minority students and lower SES students participating in the program. Using statistics tables and a power value of
Table 7

Summary of Subsidiary Questions and Statistical Analyses Associated with Research Question Three

Subsidiary Question 3-A: How do the DBA mathematics posttest scores, TerraNova mathematics national percentile ranking, and NJ ASK mathematics scaled scores of non-Asian minority Project Excel participants compare with those of White non-participants?

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Groups Compared</th>
<th>Data Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBA Posttest</td>
<td>Non-Asian Minority Participants ($n = 24$)</td>
<td>Independent</td>
</tr>
<tr>
<td>2nd &amp; 3rd Grade</td>
<td>White Non-participants ($n = 65$)</td>
<td>samples $t$ test</td>
</tr>
<tr>
<td>TerraNova</td>
<td>Non-Asian Minority Participants ($n = 28$)</td>
<td>Independent</td>
</tr>
<tr>
<td>2nd &amp; 5th Grade</td>
<td>White Non-participants ($n = 60$)</td>
<td>samples $t$ test</td>
</tr>
<tr>
<td>NJ ASK</td>
<td>Non-Asian Minority Participants ($n = 25$)</td>
<td>Independent</td>
</tr>
<tr>
<td>3rd &amp; 4th Grade</td>
<td>White Non-participants ($n = 51$)</td>
<td>samples $t$ test</td>
</tr>
</tbody>
</table>
Subsidiary Question 3-B: How do the DBA mathematics posttest scores, TerraNova mathematics national percentile ranking, and NJ ASK mathematics scaled scores of non-Asian minority Project Excel participants compare with the district, state, and national norms?

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Groups Compared</th>
<th>Data Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBA Posttest</td>
<td>Non-Asian Minority Participants ($n = 24$)</td>
<td>One-sample $t$ test</td>
</tr>
<tr>
<td>2nd &amp; 3rd Grade</td>
<td>District DBA Posttest Average</td>
<td></td>
</tr>
<tr>
<td>TerraNova</td>
<td>Non-Asian Minority Participants ($n = 28$)</td>
<td>One-sample $t$ test</td>
</tr>
<tr>
<td>2nd &amp; 5th Grade</td>
<td>TerraNova National Percentile Norm</td>
<td></td>
</tr>
<tr>
<td>NJ ASK</td>
<td>Non-Asian Minority Participants ($n = 25$)</td>
<td>One-sample $t$ test</td>
</tr>
<tr>
<td>3rd &amp; 4th Grade</td>
<td>NJ ASK State Average</td>
<td></td>
</tr>
</tbody>
</table>
Table 7 (continued)

Subsidiary Question 3-C: How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scaled scores of lower SES participants compare with those of moderate/higher SES non-participants?

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Groups Compared</th>
<th>Data Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBA Posttest</td>
<td>Lower SES Participants ($n = 27$)</td>
<td>Independent</td>
</tr>
<tr>
<td>2nd &amp; 3rd Grade</td>
<td>Mod./Higher-SES Non-participants ($n = 64$)</td>
<td>samples $t$ test</td>
</tr>
<tr>
<td>TerraNova</td>
<td>Lower SES Participants ($n = 23$)</td>
<td>Independent</td>
</tr>
<tr>
<td>2nd &amp; 5th Grade</td>
<td>Mod./Higher-SES Non-participants ($n = 76$)</td>
<td>samples $t$ test</td>
</tr>
<tr>
<td>NJ ASK</td>
<td>Lower SES Participants ($n = 33$)</td>
<td>Independent</td>
</tr>
<tr>
<td>3rd &amp; 4th Grade</td>
<td>Mod./Higher-SES Non-participants ($n = 70$)</td>
<td>samples $t$ test</td>
</tr>
</tbody>
</table>
Table 7 (continued)

Subsidiary Question 3-D: How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scaled scores of lower SES participants compare with the district, state, and national norms?

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Groups Compared</th>
<th>Data Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBA Posttest</td>
<td>Lower SES Participants (n = 27)</td>
<td>One-sample t test</td>
</tr>
<tr>
<td>2nd &amp; 3rd Grade</td>
<td>District DBA Posttest Average</td>
<td></td>
</tr>
<tr>
<td>TerraNova</td>
<td>Lower SES Participants (n = 23)</td>
<td>One-sample t test</td>
</tr>
<tr>
<td>2nd &amp; 5th Grade</td>
<td>TerraNova National Percentile Norm.</td>
<td></td>
</tr>
<tr>
<td>NJ ASK</td>
<td>Lower SES Participants (n = 33)</td>
<td>One-sample t test</td>
</tr>
<tr>
<td>3rd &amp; 4th Grade</td>
<td>NJ ASK State Average</td>
<td></td>
</tr>
</tbody>
</table>
0.20, a significance level of 0.05 and a large effect size of 1.00, the appropriate minimum sample size was determined to be 17 in each group (Hinkle, Wiersma & Jurs, 2003, p. 654). District policy required that teachers document student and teacher attendance throughout the extended-day program, the class size, the duration of the program, the number of hours each teacher spent receiving training and participating in professional development related to the program, and the amount of instructional time spent with each student. Data were used to account for variability in the experiences each child had and implementation fidelity. Written records were examined for patterns in learning goals taught, instructional methodologies used, curriculum materials, student and teacher attendance, and time on task. The researcher examined student portfolios and coded teacher anecdotal notes based on references to math standards taught and key words such as addition, volume, probability, tree diagrams, and patterns. These key words allowed the researcher to group content taught based on the learning goals and the NJ math standards.

Hypotheses

Although trends in research characterized the existence of achievement gaps between non-Asian minorities and White students and between students of different SES, the researcher could not assume that similar trends could be found when examining the math achievement for the population of Riverview students. Hence the researcher conducted an investigation to answer the following question: What is the significance, if any, of preexisting differences in mathematical abilities of the participating groups in this study? District leaders anticipated that there would be significant pretreatment
differences in DBA pretest scores prior to participation in the Project Excel program between the following groups: (a) Project Excel participants and non-participants, (b) non-Asian minority students and White students, and (c) lower SES students and moderate/higher SES students, and (d) no significant difference in DBA pretest scores between the Project Excel participants and the Title I-eligible non-participants. By examining the relationship between these groups the researcher aimed to gain insight as to the baseline performance of the groups when the treatment was not given.

The null hypotheses for this investigation suggested there would be no differences in pretest scores for the various pairwise combinations mentioned previously. The null hypotheses were expressed as follows:

\[ H_0: \mu_A = \mu_B = \mu_C = \mu_D = 0 \]

The alternate hypotheses suggested that there would be differences in the pretest scores for the four pairwise combinations. The alternate hypotheses were expressed as follows:

\[ H_A: \mu_A \neq \mu_B \neq \mu_C \neq \mu_D \neq 0 \]

In the second investigation the researcher answered the following question: What is the relationship between participation in Title I programs, specifically the Project Excel program, and mathematics achievement? Test data collected from Title I-eligible students who participated in the program (Group A) were compared to the test data collected from three reference groups: (a) all non-participants (Groups B & C), (b) district, state, and national norms, and (c) Title I-eligible non-participants (Group B).

The null hypotheses for these investigations suggested that there would be no differences in DBA mathematics posttest scores, TerraNova mathematics percentile rankings, and NJ
ASK mathematics scores between the groups. The null hypothesis was expressed as follows:

\[ H_0: \mu_{AC} = \mu_{HC} = 0 \quad H_0: \mu_{AC} = \mu_{norm} = 0 \quad H_0: \mu_{AC} = \mu_{EE} = 0 \]

The alternate hypothesis suggested that there would be differences in DBA mathematics posttest scores, TerraNova mathematics mean percentile rankings, and NJ ASK mathematics scores between the groups. The alternate hypothesis was expressed as follows:

\[ H_A: \mu_{AC} - \mu_{HC} \neq 0 \quad H_A: \mu_{AC} - \mu_{norm} \neq 0 \quad H_A: \mu_{AC} - \mu_{EE} \neq 0 \]

In the third investigation the researcher answered the following research question:

What is the relationship between participation in Title I programs, specifically the Project Excel program and closing the minority and SES achievement gap? The researcher compared the performance of non-Asian minority students who participated in the program with that of two reference groups: (a) White students who did not participate in the programs and (b) district, state, and national norms. The null hypothesis for this investigation suggested that there would be no differences in DBA mathematics posttest scores, TerraNova mathematics percentile rankings, and NJ ASK mathematics scores between the groups. The null hypothesis was expressed as follows:

\[ H_0: \mu_{AC} > 0 \quad H_0: \mu_{AC} = \mu_{norm} = 0 \]

The alternate hypothesis suggested that there would be differences in DBA mathematics posttest scores, TerraNova mathematics percentile rankings, and NJ ASK mathematics scores between the groups. The alternate hypothesis was expressed as follows:

\[ H_A: \mu_{AC} - \mu_{EE} \neq 0 \quad H_A: \mu_{AC} - \mu_{norm} \neq 0 \]
To address the SES achievement gap, the researcher compared the performance of lower SES students who participated in the program with that of two reference groups: (a) moderate/higher SES students who did not participate in the program and (b) district, state, and national norms. The null hypothesis for this investigation suggested that there would be no differences in DBA mathematics posttest scores, TerraNova mathematics percentile rankings, and NJ ASK mathematics scores between the groups. The null hypothesis was expressed as follows:

\[ H_0: \mu_{low} - \mu_{mod/high} = 0 \quad H_1: \mu_{low} - \mu_{mod/high} = 0 \]

The alternate hypothesis suggested that there would be differences in DBA mathematics posttest scores, TerraNova mathematics percentile rankings, and NJ ASK mathematics scores between the groups. The alternate hypothesis was expressed as follows:

\[ H_0: \mu_{low} - \mu_{mod/high} \neq 0 \quad H_1: \mu_{low} - \mu_{mod/high} \neq 0 \]

Summary

Riverview district leaders had a vested interest in the outcomes of this investigation. Leaders were hopeful that the study would begin to provide statistical evidence to reinforce their perceived notions that participation in the extended-day programs did increase mathematics achievement of Title I-eligible students and helped to close the achievement gap.

Chapter 3 provided a description of the research design, participants, and instrumentation to be used in the present study. Details were also provided on the type
and structure of the extended-day program, the method for data collection and analysis, hypotheses to be tested, and variables to be examined in the present study.

Chapter 4 provides a detailed statistical analysis of the data, a summary of the results of the investigation, and an interpretation of the descriptive and inferential findings.
Chapter IV

Data and Analyses

Introduction

Ladson-Billings (1994) remarked:

No challenge has been more daunting than that of improving the academic achievement of African American students. Burdened with a history that includes the denial of education, separate and unequal education, and relegation to unsafe, substandard inner-city schools, the quest for quality education remains an elusive dream for the African American community. However, it does remain a dream—perhaps the most powerful for the people of African descent in this nation. (p. ix)

Over 10 years later was the quest still just a dream or did the political prowess of the nation’s leaders find a solution in NCLB? Billions of dollars in Title I funds have been allocated to schools and districts for the purpose of making the dream a reality, yet has the outpouring of federal funds made a difference?

In the present study the researcher described the relationship between participation in a Title I-funded extended-day mathematics program, Project Excel, and academic achievement in mathematics. To investigate the nature of the relationship the researcher addressed the following research questions:

1. What is the significance, if any, of preexisting differences in mathematical abilities of the participating groups in this study?
2. What is the relationship between participation in Title I programs, specifically the Project Excel program, and mathematics achievement?

3. What is the relationship between participation in Title I programs, specifically the Project Excel program, and closing the achievement gaps between low-income and moderate/high-income students and non-Asian minority and White students?

Archival data were gathered on 406 students in Grades 2 through 5 and coded by a confidential secretary to preserve student anonymity. After excluding students who received additional mathematics instruction as a part of their regular school day, 284 students were included in this study. To address the first research question, 162 students in Grades 2 through 5 were invited to participate in an extended-day program designed to provide them with an additional 60 to 75 minutes of after school mathematics instruction weekly. The building principal and a committee of teachers selected these students based on free and reduced lunch status, a proxy for poverty or SES, previous year’s standardized test scores, recommendations from their teachers, and their performance on a grade-level DBA pretest in mathematics. Of the 162 Title I-eligible students, 87 students opted to participate. Participation was voluntary and required parental permission. The 87 participating students were labeled Group A, the group receiving the treatment. The remaining 75 students who were eligible but elected not to participate were labeled Group B, one of the control or comparison groups. The remaining 122 participants, most of whom were children from moderate/high income families and were ineligible to receive Title I services, were labeled Group C. Although all students took the DBA mathematics pretest during the first week of school, only the second- and third-
grade students took the DBA mathematics posttest in June. Second- and fifth-grade students took the TerraNova test, and third- and fourth-grade students took the NJ ASK test in the spring. Due to the variety of spring assessments administered at each grade level, the researcher partitioned the three groups to allow for comparison of math performance based on which spring assessment (DBA posttest, TerraNova, or NJ ASK) was taken.

Student performance on the mathematics pre- and posttests was reported as percent items answered correctly on a scale of 0 to 100%. Scores for students taking the TerraNova test were reported as national percentile rankings on a scale of 1st to 99th percentile. The NJ ASK raw scores were converted into scaled math scores that ranged from 0 to 300 points. An analysis of student performance sheets, attendance records, and time sheets provided descriptive information regarding the number and type of learning objectives taught in the program; the type of assessment used in the program; the allotted time for direct instruction, computer-assisted instruction, and assessment; the curriculum used; and the instructional strategies incorporated.

The growing pressures to meet AYP, the tight regulations imposed by NCLB and the scarcity of resources led district leaders to question whether the Project Excel program was a worthwhile investment for improving academic achievement and closing the minority and SES achievement gaps. District leaders anticipated that Project Excel participants (Group A) would demonstrate significantly greater mathematics achievement on the spring assessments than would the Title I-eligible non-participants (Group B). District leaders hoped to see reduction in the non-Asian minority and SES achievement gaps following participation in the program. The present study served to aid district
leaders in their understanding of the relationship between participation in the program and academic achievement. Data and analyses helped the researcher to answer three main research questions and several subsidiary questions.

**Research Question One**

In the present study the researcher examined the relationship between participation in Project Excel and mathematical achievement and closing achievements gaps; however, one can hardly close a gap without first identifying its existence. Hence in this investigation the researcher determined if significant differences in mathematical performance prior to participating in Project Excel existed between students. To determine the existence of achievement gaps, the researcher compared the DRA pretest scores of four groups: (a) participants and non-participants, (b) participants, eligible non-participants, (c) lower SES participants and moderate/higher SES non-participants, and (d) non-Asian minority participants and White non-participants.

As seen in Figure 5, the mean DRA pretest score for Project Excel participants ($\mu_A = 38\%$) was 16 percentage points lower than the mean pretest score for all non-participants ($\mu_C = 54\%$), and this difference was statistically significant ($t = -6.648, p = .000$). The participants ($\mu_A = 38\%$) scored significantly lower than did the eligible non-participants ($\mu_B = 46\%, mean\ diff. = -8\%, t = -2.872, p = .005$), and than did ineligible non-participants ($\mu_C = 59\%, mean\ diff. = -21\%, t = -3.614, p = .000$). These findings established a baseline of preexisting differences in mathematical abilities between the groups prior to participation in Project Excel.
Figure 5. Mean DLA mathematics pretest scores for Groups A, B, B and C, and C. Data were collected from students who participated in Project Excel (Group A, n = 87), students who were eligible but opted not to participate (Group B, n = 75), students who were Title I-eligible and did not participate (Group C, n = 122), and all non-participants (Groups B & C, n = 197). The district mean for all students who took the DLA pretest was 49% items answered correctly. All mean differences between groups were significant (p < .05).
Did Riverview student mathematics performance reflect the same minority and SES achievement gaps found nationally? Non-Asian minority participants ($\mu_M = 35\%$) scored significantly lower on the DBA pretest than did White non-participants ($\mu_W = 56\%$, mean diff. = -21\%, $t = -7.372$, $p = .000$). Similarly lower SES participants ($\mu_{low} = 35\%$) scored significantly lower on the DBA pretest than did moderate/higher SES non-participants ($\mu_{high} = 57\%$, mean diff. = -22, $t = -7.926$, $p = .000$), thereby confirming the existence of significant pretreatment differences (Figure 5).

Summary of Research Question One

Results from these four comparisons showed that Project Excel participants scored significantly lower than did the various comparative non-participant groups on the DBA pretest. These data provide evidence of preexisting achievement gaps between participants and non-participants.

Research Question Two

The second research question asked about the relationship between participation in the Project Excel program and academic achievement in mathematics. Students in Group A were identified as Title 1-eligible and participated in an extended-day mathematics program where they received 60 to 75 minutes of additional mathematics instruction one day a week for 21 to 23 weeks. Students in Group B were also identified as Title 1-eligible but opted not to participate in the extended-day mathematics program. Group C students were not eligible, so Groups B and C comprised the non-participants.
Figure 6. Mean DBA mathematics pretest scores categorized by race and SES. The district mean for all students who took the DBA pretest was 49% items answered correctly. The mean difference between non-Asian minority participants and White non-participants and the mean difference between lower SES participants and moderate/higher SES non-participants are significant ($p < .05$).
Comparison of Project Excel participants and all Non-participants. During the 2004 – 2005 school year, district leaders required that all elementary school teachers assess their students in mathematics using a series of DBAs. These assessments were designed to measure the growth and progress students made toward attainment of the mathematics learning goals. The assessments included a pretest given during the first week of school, a posttest given in June, and nine monthly assessments. To determine the relationship between participation in Project Excel and improving academic achievement, the researcher examined the mathematics performance on the DBA mathematics posttest, the TerraNova mathematics test, and the NJ ASK mathematics test between the participants (Group A) and the non-participants (Groups B and C). Project Excel second- and third-grade participants (Group A) who took the DBA mathematics posttest answered an average of 63% of items correctly; by comparison all non-participants (Groups B and C) answered an average of 78% items correctly (Figure 7). The mean difference (-15%) was statistically significant ($t = -5.867$, $p = .000$) and indicated that non-participants scored significantly higher than did Project Excel participants.

Second- and fifth-grade Project Excel participants who took the TerraNova mathematics test had an average national ranking of the 59th percentile; by comparison all non-participants had an average national ranking of the 71st percentile (Figure 8). The mean difference in percentile rankings (12 NCEs) was statistically significant ($t = -3.172$, $p = .002$) and suggested that non-participants scored significantly higher than did Project Excel participants.
Figure 7. Mean DBA mathematics posttest scores for participants and non-participants. Data were collected from second- and third-grade students who participated in Project Excel (n = 35) and those who did not (n = 113). Posttest scores were the percent items answered correctly. Mean difference (-15%) between participants and non-participants was statistically significant (t = -5.807, p = .000). The district mean for all students who took the DBA posttest was 75% items answered correctly.
Figure 8. Mean TerraNova mathematics percentile rankings for participants and non-participants. Data collected were from second- and fifth-grade students who participated in Project Excel \( (n = 41) \) and those who did not \( (n = 101) \). TerraNova scores were reposted as national percentile rankings. The mean difference \(-12 \text{ NCEs}\) between participants and non-participants was statistically significant \( t = -3.172, p = .002 \). The national mean for all students who took the TerraNova was 50th percentile.
Third and fourth-grade Project Excel participants had an average NJ ASK scaled mathematics score of 204; by comparison all non-participants had an average scaled mathematics score of 234 (Figure 9). The mean difference of -30 scaled points was statistically significant ($t = -6.369$, $p = .000$) and indicated that non-participants scored significantly higher than did Project Excel participants. An analysis of data collected from the three spring assessments comparing the participant/non-participant performance showed that non-participants performed better on the DIBA posttest, TerraNova, and NJ ASK than did Project Excel participants.

Comparison of Project Excel participants to district, state, and national norms.

In the previous section, the researcher outlined how the mathematics performance of the participants compared with that of the non-participants; however, examination of the data on a more global level led the researcher to question how did the mathematics performance participants compare to district, state, and national norms on the three spring assessments. To address that question, the researcher compared the performance of the Project Excel participants to the district average on the DIBA posttest, the national percentile ranking average on the TerraNova, and the state average on the NJ ASK. Second- and third-grade participants ($\mu = 63\%$) scored significantly lower than the DIBA posttest average of 75% items answered correctly (mean diff. = -12%, $t = -5.042$, $p = .000$). Third and fourth-grade participants ($\mu = 204$) scored significantly lower on the NJ ASK mathematics assessment than the state average scaled mathematics score of 228 (mean diff. = -24 scaled points, $t = -6.115$, $p = .000$). Examination of DIBA posttest scores and NJ ASK scores suggested that non-participants performed better than did Project Excel participants.
Figure 9. Mean NJ ASK mathematics scores for participants and non-participants. Data were collected from third- and fourth-grade students who participated in Project Excel (n = 46) and those who did not (n = 96). NJ ASK scores were reported as scaled scores from 0 – 300. The mean difference (-30 scaled points) between participants and non-participants was statistically significant (t = -6.369, p = .000). The state mean for all students who took the NJ ASK 3 and 4 math assessments was 228.
Second- and fifth-grade participants ($\mu = 59^{th}$ percentile) scored significantly higher on the TerraNova mathematics test than the national mean of $50^{th}$ percentile ($mean \ diff. = 9$ NCEs, $t = 2.671, p = .011$). This was evidence to support the claim that participants performed significantly better on the TerraNova than did other students around the country.

The results of the comparison of the mathematics performance of Project Excel participants to district, state, and national norms yielded inconclusive findings. Results from the TerraNova data show that Project Excel students performed significantly better than the norm; yet these results were not supported by evidence from the DBA posttest or the NJ ASK data. The researcher cautions district leaders in their interpretation of these findings due to the low reliability and high SEM reflected in the NJ ASK tests.

**Comparison of Project Excel participants and Title I-eligible non-participants.**

The comparison of the performance of Project Excel students (Group A) to all non-participants and district, state, and national norms provided useful information, however the researcher questioned how the performance of Project Excel students compared to that of other students with similar mathematics backgrounds and SES status, the Title I-eligible non-participants (Group B). After a year of participating in the Project Excel program second and third-grade students ($\mu = 63\%$) answered fewer questions correctly on the DBA posttest than did the Title I-eligible students ($\mu = 71\%$) who opted not to participate in the program. The mean difference (-8%) was statistically significant ($t = -2.253, p = .027$) as shown in Figure 10.
Figure 10. Mean DBA mathematics posttest scores for participants and eligible non-participants. Data were collected from second- and third-grade Title I-eligible students who participated in Project Excel (n = 35) and Title I-eligible students who did not (n = 39). Posttest scores were the percent items answered correctly. Mean difference (8%) between participants and eligible non-participants was statistically significant (t = -2.253, p = .027). The district mean for all students who took the DBA posttest was 75% items answered correctly.
Figure 11 shows that second- and fifth-grade Project Excel students (Group A) who took the TerraNova mathematics test had a mean mathematics ranking of 59th percentile. Although these students had a higher average percentile ranking than the Title I-eligible non-participants’ (Group B) mean ranking of 54th percentile, the mean difference between the two groups (5 NCEs) was not statistically significant ($t = .970, p = .335$).

Similar results were reflected in the NJ ASK scaled scores for third- and fourth-grade students (Figure 12). There was no significant difference between the math performance of the participants (Group A) and the eligible non-participants (Group B) who had mean scaled mathematics scores of 204 and 212, respectively ($mean\ diff. = -8, t = -1.295, p = .199$). The results of the comparison of the performance of Project Excel participants to that of eligible non-participants yielded inconclusive findings. There were no statistically significant differences between the performance of Project Excel participants and the eligible non-participants on the TerraNova test and the NJ ASK, suggesting that Project Excel participants performed just as well as did eligible non-participants. DBA posttest results reflected statistically significant differences that indicated that Project Excel participants did not perform as well as did eligible non-participants and in fact performed worse. Further investigation was needed to infer whether a relationship existed between participation in Project Excel and improved academic achievement.
Figure 11. Mean mathematics TerraNova percentile rankings for participants and eligible non-participants. Data were collected from second- and fifth-grade Title I-eligible students who participated in Project Excel (n = 41) and Title I-eligible students who did not (n = 37). TerraNova scores were reported as national percentile rankings. Mean difference (5 NCEs) between participants and eligible non-participants were not statistically significant (t = -0.970, p = 0.335). The national mean for all students who took the TerraNova was 50th percentile.
Figure 12. Mean mathematics NJ ASK scores for participants and eligible non-participants. Data were collected from third- and fourth-grade Title I-eligible students who participated in Project Excel (n = 46) and Title I-eligible students who did not (n = 38). NJ ASK scores were reported as scaled scores from 0 to 300. Mean difference (.8 scaled points) between participants and eligible non-participants was NOT statistically significant (t = 1.295, p = .199). The state mean for all students who took the NJ ASK 3 and 4 math assessments was 228.
Summary of Research Question Two

The researcher proposed three subsidiary questions as a part of the second investigation: (a) How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scores of Project Excel participants compare with those of all non-participants; (b) How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scores of Project Excel participants compare with the district, state, and national norms; and (c) How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scores of Project Excel participants compare with those of the Title I-eligible non-participants?

The researcher found that the Project Excel students scored significantly lower on the DBA posttest, the TerraNova and the NJ ASK than did all non-participants. This negative correlation suggested that participation may actually be harmful for children. When compared to the eligible non-participants, Project Excel students showed no significant difference in performance on TerraNova and NJ ASK and performed significantly lower on the DBA posttest than did the eligible non-participants. The most promising results came from the TerraNova norm study. Project Excel participants performed significantly better than the TerraNova norm, yet performed significantly lower than the district DBA posttest average and the state average on the NJ ASK. The implications of these findings for district leaders are presented in chapter 5.
Research Question Three

NCLB has mandated a system of tiered penalties for schools and districts that fail to meet AYP, for a given subgroup several years in a row. These penalties include intra- and inter-district transfers, mandatory supplemental educational services, reassignment of staff and state take-over. Often it is the minority and lower SES populations that fail to meet AYP and one of the biggest challenges facing district leaders was how to close the minority and SES achievement gaps. Black and Hispanic students comprised more than 60% of the Project Excel participant population and over 64% of the participants qualified to receive free or reduced lunch. In the second investigation the researcher addressed the following research question: What is the relationship between participation in Title I programs, specifically in Project Excel and closing the achievement gaps between low-income and moderate/high-income students and non-Asian minority and White students?

Minority achievement gap. To address the minority achievement gap, the researcher compared the academic performance of non-Asian minority Project Excel participants to two reference groups: (a) White non-participants and (b) district, state, and national norms. Second- and third-grade Project Excel non-Asian minority participants who took the DBA posttest answered an average of 60% of the items correctly, yet scored 20 percentage points lower than did White non-participants who answered an average of 80% of the items correctly. The mean difference between the groups was statistically significant (t = -6.429, p = .000). Figure 13 shows that when participant performance is compared to the district norm of 75% items answered correctly on the DBA posttest, non-Asian minority participants (μ = 60%) performed significantly lower than the norm.
Figure 13. Mean DBA mathematics posttest scores for non-Asian minority and White non-participants. Data were collected from second- and third-grade students. Posttest scores were the percent items answered correctly. Mean difference (-20%) between non-Asian participants and White non-participants was statistically significant ($t = -6.429, p = .000$). Mean difference (-22%) between lower SES participants and moderate/higher SES non-participants was statistically significant ($t = -8.528, p = .000$). The mean for all students who took the posttest was 75% items answered correctly.
(mean diff. = -15%, t = -5.639, p = .000). The DBA posttest study showed that there was a negative relationship. Non-Asian minority participants performed significantly lower on the DBA posttest than did White non-participants and lower than the norm.

Project Excel non-Asian minority participants who took the TerraNova had a mean national ranking of 55th percentile; by comparison White non-participants had a mean national ranking of 71st percentile (Figure 14). The difference between the groups (-18 NCEs) was statistically significant (t = -4.044, p = .000). There was no statistically significant difference in the performance of non-Asian minority participants and the national TerraNova norm of 50th percentile (mean diff. = 5 NCEs, t = 1.152, p = .259).

Although they did not perform as well as the White non-participants, non-Asian minority students performed just as well as the average student across the country who took the TerraNova.

Non-Asian minority participants earned an average NJ ASK scaled mathematics score of 196; by comparison White non-participants earned an average scaled mathematics score of 241 (Figure 15). The mean difference between the groups (45 scaled math points) was statistically significant (t = -8.592, p = .000). Non-Asian minority participants scored significantly lower than the NJ ASK state average of 228 (mean diff. = 32 scaled points, t = -6.680, p = .000). The NJ ASK results were similar to the DBA results – a negative relationship existed. Non-Asian minority students participating in the program performed significantly lower on the NJ ASK than did White non-participants and the state norm.

Although Riverview district leaders anticipated that participation in the Project Excel program would lead to gains in closing the minority achievement gap, the results
Figure 14. Mean TerraNova mathematics percentile rankings for non-Asian minority and White non-participants. Data were collected from second- and fifth-grade students. TerraNova scores were reported as national percentile rankings. Mean difference (-18 NCEs) between non-Asian participants and White non-participants was statistically significant ($t = -4.044$ and $p = .000$). Mean difference (-23 NCEs) between lower-SES participants and moderate/higher SES non-participants was statistically significant ($t = -6.069$, $p = .000$). The mean for all students who took the TerraNova nationally was 50th percentile.
Figure 15. Mean NJ ASK mathematics scores for non-Asian minority and White non-participants. Data were collected from third- and fourth-grade students. NJ ASK scores were reported as scaled scores from 0 to 300. Mean difference (45 scaled points) between non-Asian participants and White non-participants was statistically significant ($t = -8.592, p = .000$). Mean difference (41 scaled points) between lower SES participants and moderate/higher SES non-participants was statistically significant ($t = -9.024, p = .000$). The state mean for all students who took the NJ ASK 3 and 4 math assessments was 228.
from this investigation suggested that no significant gains were made, and at best, non-Asian minority students performed the same as the average student who took the TerraNova.

**SES achievement gap.** One of the drawbacks of the NCLB subgroup classification system was that most students fell under multiple categories and a low-income, minority, special education student can contribute to a district failing three separate categories. It was no surprise that those of the minority achievement-gap investigation. Lower SES students, those receiving free or reduced lunch, answered 59% of the DBA posttest items correctly; by comparison moderate/higher SES non-participants answered 81% of the posttest items correctly (Figure 1J). The mean difference of -2.2 percentage points was statistically significant ($t = -8.528, p = .000$). When their performance was compared to the district average of 75% items answered correctly, lower SES participants scored significantly lower than the average (*mean diff. $\approx -16\%$, $t = -6.453, p = .000$). Based on the findings from the DBA posttest, lower SES participants did not show any evidence of closing the achievement gap on the DBA posttest.

Lower SES Project Excel participants' average national ranking on the TerraNova was 52nd percentile; by comparison the moderate/higher SES non-participants had an average national ranking of 75th percentile (Figure 14). The mean difference (-23 NCEs) was statistically significant ($t = -4.936, p = .000$). There was no statistically significant difference between lower SES participant TerraNova performance and the national norm of 50th percentile ($t = .428, p = .673$). These results reflect the same trends that the non-Asian minority students experienced. Although the findings provided no evidence of gap
shrinkage, they did show that lower SES participants performed equally well as the average student who took the TerraNova.

Lower SES participants earned an average NJ ASK scaled mathematics score of 201; by comparison the moderate/higher SES non-participants earned an average scaled mathematics score of 242 (Figure 15). The mean difference of -41 scaled points was statistically significant (t = -9.02, p = .000). These students scored significantly lower than the state average of 228 (mean diff. = -27 scaled points, t = -5.875, p = .000). The many minority students who participated in the Project Excel program were also from low-income families and that the results from the SES achievement-gap investigation mirrored findings from the NJ ASK investigation resulted in a negative relationship and implied that participation in the program may lead to students performing worse than their peers.

Although one Project Excel program goal was to close the achievement gap between the different SES classes, results from this investigation were inconclusive as to the relationship that participation in the program had on achieving that goal. Data from the NJ ASK and DBA posttests indicated lack of evidence in support of achieving that goal; however, data from the TerraNova indicated that lower SES students performed as well as the national norm.

Summary of Research Question Three

The researcher proposed four subsidiary questions affiliated with the third research question regarding closing the achievement gaps:
3-A How do the DBA mathematics posttest scores, TerraNova mathematics national percentile ranking, and NJ ASK mathematics scaled scores of non-Asian minority Project Excel participants compare with those of White non-participants?

3-B How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scaled scores of non-Asian minority Project Excel participants compare with the district, state, and national norms?

3-C How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scaled scores of lower SES participants compare with those of moderate/higher SES non-participants?

3-D How do the DBA mathematics posttest scores, TerraNova mathematics national percentile rankings, and NJ ASK mathematics scaled scores of lower SES participants compare with the district, state, and national norms?

The findings from the minority and SES investigations provided no clear evidence that non-Asian minority and lower SES Project Excel participants performed significantly better than did White and moderate/higher SES non-participants. There was evidence that non-Asian minority and lower SES students performed at least as well as the national norm on the TerraNova mathematics test.

The researcher observed that the findings from the first two investigations revealed several trends in performance. The DBA posttest scores were used as an indicator of improved student achievement for seven comparisons of Project Excel participant performance to that of various groups of non-participants and district norms.
In all seven comparisons, the Project Excel students scored significantly lower than the
district norms and significantly lower than did the various non-participating groups.

A similar trend was observed when the NJ ASK scores were used as the indicator
of improved academic achievement. Project Excel participants performed significantly
lower than the state norms and significantly lower than did three of the four non-
participating comparison groups. There was no significant difference in NJ ASK
mathematics performance between the Project Excel participants and the Title I-eligible
non-participants.

The TerraNova results suggested a different trend in performance. Project Excel
participants scored significantly lower than did all non-participants, White non-
participants and moderate/higher SES non-participants. However, there was no
significant difference in performance of non-Asian minority participants and the national
norm, lower SES participants and the national norm, and the participants and that of the
Title I-eligible non-participants. Project Excel participants scored significantly higher
than the national norm. The potential implications of these trends in performance are
explored in more detail in chapter 5.

District leaders should use caution when interpreting these findings and making
conclusions regarding the effectiveness of the Project Excel program. Figures 16 and 17
show the pre- and posttest gain scores for the various groups participating in this study.
Project Excel participants earned higher gains than did all other non-participating groups.
Although these higher gains were not statistically significant, the scores did indicate that
Project Excel participants increased their mathematical performance at the same rate as or
Figure 16. Mean DBA pre-and posttest gains for Groups A, B, B and C, and C. Data were collected from second- and third-grade students who took both assessments. Project Excel participants increased their performance at a similar level as all other non-participants. Although Project Excel students had the highest gain, the differences in the mean performance between the groups were not statistically significant ($p > .05$).
Figure 17. Mean DBA pre- and posttest gains categorized by race and SES. Data collected from second- and third-grade students who took both assessments. Non-Asian minority and lower SES Project Excel participants increased their performance at a similar level as all other non-participants. Although Project Excel students had the higher gains, the differences in the mean performance between the groups were not statistically significant ($p > .05$).
at a greater rate than other non-participants (Table 8). The potential implications of this finding for district policy and practice are discussed in chapter 5.

Table 8  

**DBA Pre- and posttest Gains for Second and Third-Grade Students**

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>μ</th>
<th>Diff.</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Excel participants</td>
<td>35</td>
<td>23.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All non-participants</td>
<td>113</td>
<td>21.37</td>
<td>2.057</td>
<td>.640</td>
<td>.523</td>
</tr>
<tr>
<td>Eligible non-participants</td>
<td>39</td>
<td>20.54</td>
<td>2.890</td>
<td>.751</td>
<td>.455</td>
</tr>
<tr>
<td>Ineligible non-participants</td>
<td>74</td>
<td>21.81</td>
<td>1.618</td>
<td>.471</td>
<td>.635</td>
</tr>
<tr>
<td>Non-Asian minority</td>
<td>24</td>
<td>22.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White non-participants</td>
<td>65</td>
<td>21.26</td>
<td>1.330</td>
<td>.398</td>
<td>.692</td>
</tr>
<tr>
<td>Lower SES participants</td>
<td>27</td>
<td>22.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Higher-SES non-participants</td>
<td>84</td>
<td>20.95</td>
<td>2.011</td>
<td>.558</td>
<td>.578</td>
</tr>
</tbody>
</table>
Additional Findings

The previous section outlined the results from the quantitative analysis of performance data collected on the 284 students participating in the study. Additional quantitative and qualitative data were collected to provide information regarding student and teacher attendance rates, class size, usage of time, and instructional practices used in the Project Excel program.

Class size and attendance. Achilles and Finn (2001) reported the benefits of small class size participation. District leaders aimed to keep the Project Excel class size small to afford the maximum student-teacher interaction feasible. The average class size for second grade was 3.5 students, third grade was 4.4 students, fourth grade was 3.5 students, fifth grade was 3.7 students, and overall was 3.7 students. Students who failed to attend the program on a regular basis were dropped from the program and did not participate in the study. Records indicated that only five students were dropped from the program during the first few weeks.

The student attendance rates were 92% for Grade two, 72% for Grade three, 74% for Grade four, 75% for Grade five, and 77% overall. The teacher attendance rate was 96%. The second-grade program was held at the primary school. Second-grade teachers regularly walked their students to their Project Excel classroom at the end of the day, thereby assuring their attendance. At the elementary school, third-, fourth- and fifth-grade students had 5 minutes of unsupervised time between the end of the school day and the start of the program. Students were not escorted to the program by their classroom teachers, and several students walked home instead of attending. The disparities in student attendance may have been attributed to the program location.
Time, methodology, and curriculum. Students in Grade two were provided 60 minutes of weekly instruction, and students in Grades three through five were provided 75 minutes of weekly instruction. The individual student portfolios provided details regarding the use of this time and how students were instructed, what materials they used, and what learning goals they attended to each day. Table 9 categorizes the percent of instructional time spent engaging in direct/hands-on instruction, computer-assisted instruction, and assessment.

Table 9

<table>
<thead>
<tr>
<th>Grade</th>
<th>Direct/hands-on</th>
<th>Computer-aided</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>80.1%</td>
<td>0.9%</td>
<td>19.0%</td>
</tr>
<tr>
<td>3</td>
<td>69.8%</td>
<td>3.5%</td>
<td>26.3%</td>
</tr>
<tr>
<td>4</td>
<td>77.9%</td>
<td>1.7%</td>
<td>20.4%</td>
</tr>
<tr>
<td>5</td>
<td>70.9%</td>
<td>11.7%</td>
<td>17.4%</td>
</tr>
<tr>
<td>Average</td>
<td>74.7%</td>
<td>4.8%</td>
<td>20.5%</td>
</tr>
</tbody>
</table>

Although the curriculum used in the Project Excel program varied from grade to grade, all curricula were standards-based. Anecdotal notes regarding lesson objectives written by instructors were examined and coded based on the NJ math standard they matched. In most cases the researcher relied on key words to classify instructional focus as (a) Number Sense and Numerical Operations, (b) Geometry and Measurement, (c) Patterns and Algebra, (d) Data Analysis, Probability and Discrete Mathematics, and (e)
Problem Solving. Figure 18 shows the average percent of instructional time dedicated to teaching each of the five clusters.

**Staffing and staff training.** Professional development was the area least mentioned in the anecdotal notes written by the Project Excel instructors. All teachers participated in two 2-hour workshops, one at the beginning of the program and one at the end of the program. The focus of these workshops was on procedures and policies rather than instructional strategies and best practices. Throughout the school year Project Excel instructors had the opportunity to partake in the professional development courses offered by the district, yet no data were collected regarding which courses were taken and the implementation of strategies learned in connection with the Project Excel program.

Project Excel teachers met monthly with each other and regular education teachers to discuss the progress the students made, yet very little data regarding the nature of those discussions was available. IRB regulations prohibited the researcher from observing the monthly discussions due to the supervisory relationship the researcher held with the instructors. Such participation may have provided the researcher with rich information regarding professional development practices, procedures, and strategies implemented by the Project Excel instructors.

IRB regulations precluded the researcher from conducting an intensive program evaluation for the purpose of ensuring implementation fidelity. All data were coded and only accessible to the researcher at the conclusion of the Project Excel program. The researcher did not conduct Project Excel classroom observations, interview participants, or coach, mentor, and evaluate teachers in connection with the program, although all of
Figure 18. Analysis of average Project Excel instructional time spent teaching various mathematical concepts. Data were obtained by coding anecdotal teacher notes written on student performance sheets of students who participated in Project Excel (n = 87).
these activities along with program evaluation, research and design fell under the realm of the job responsibilities of the researcher.

**Summary**

Chapter 4 provided a detailed summary of the comparative mathematics performance the 284 of second- through fifth-grade students who participated in the study. Due to the variability of post-treatment assessments given at different grade levels, the researcher partitioned the groups to allow for appropriate comparisons. The findings from several investigations were presented including a comparison of mathematics achievement between (a) Title I-eligible students who participated in Project Excel and Title I-eligible students who elected not to participate, (b) non-Asian minority students who participated in the program and White students who did not participate, and (c) lower SES students who participated in the program and moderate/higher SES students who did not participate.

The DBA pre-and posttest analyses revealed that second- and third-grade Project Excel students, non-Asian minority participants, and lower SES participants scored significantly lower on the pre and posttests than did the comparable groups, yet no significant differences in pre-and posttest gains were found between the groups. These findings indicated that the participating students were progressing at the same pace as all other non-participating students.

The NJ ASK analyses showed that third- and fourth-grade Project Excel participants, non-Asian minority participants and lower SES participants scored significantly lower than did the comparable groups with one exception. There was no
statistically significant difference between the mathematics performance of the participants and that of Title I-eligible non-participants.

More promising results were obtained from the TerraNova analyses. With regard to the minority and SES achievement gaps, Non-Asian minority and lower SES Project Excel participants' mathematics performance was similar to the national norm. The researcher found that Project Excel participants scored significantly higher than the national norm. This was the only case where Project Excel students performed significantly higher than did the comparison group. No significant differences were found between the mathematics performance of Project Excel participants and that of Title I-eligible non-participants. On the district level, participants scored significantly lower on the TerraNova than did all non-participants; non-Asian minority participants scored significantly lower than did White non-participants; and lower SES participants scored significantly lower than did moderate/higher SES non-participants.

A summary of findings, conclusions and interpretations of the results are presented in chapter 5. The results of this study have some practical applications for district leaders involved in policy-making and effective practice. The researcher revisits the limitations of this study and proposes ideas and other investigations for further research.
Chapter Y

Summary of Findings, Conclusions, Discussion and Recommendations

Introduction

Over 10 years ago the ASCD Improving Student Achievement Panel wrote:
The combination of poverty and diversity plays havoc with performance in
school. Disproportionately high numbers of minority, immigrant, and poor
children perform consistently in the lower third academically in U.S. schools.
The education course charted in the next few years will play a major role in
determining whether we can truly educate everybody’s children to be successful,

Today district leaders are still trying to live up to the challenge of educating everybody’s
children. International, national, state, and local data collected from TIMSS, NAEP, NJ
HSPA, NJ GEPA, NJ ASK, and DBA pre- and posttests suggest that American children
are not performing as well as children in other nations and minority and lower SES
American students are severely underperforming when compared to other students in this
nation.

Reichstein (2004a, 2004b, 2005), Barton (2004), Berliner (2005) and Hart and
Risley (2003) linked poor academic performance to several indicators of poverty, lower
SES and race. They found that the minority and SES achievement gaps were well formed
before the child entered kindergarten. In their 2 ½ year longitudinal study, Hart and
Risley concluded that by the time they reached the age of 4, children in professional
familiar would have had experiences with an average of 45 million words, children in working-class families would have had experiences with an average of 26 million words, and children in welfare families would have had experiences with an average of 13 million words. Not only did the welfare children have fewer experiences than did the professional and working-class children, they also learned new words at a slower rate than did the working-class and professional children. District leaders are left with the challenge of dealing with these first day of school inequities. The federal government allocated Title I resources to schools and districts to level the playing field between lower SES students and moderate/higher SES students. Some districts implemented schoolwide programs while others invested in extended-day programs. Riverview district leaders have provided extended-day experiences for at-risk students since 2001.

Project Excel, Riverview's extended-day mathematics program, incorporated many of the research-based opportunity-to-learn (OTL) variables that are linked to improved academic achievement. Project Excel was designed to supplement instruction, not supplant it like many earlier pull-out and basic skills programs. Students received additional time-on-task that was used to extend lessons taught during the day, reinforce basic concepts, and pre-teach future lessons. Class size was intentionally small (2-4 students) to maximize student engagement and student-teacher interaction. The curriculum was modified to meet individual student needs, and teachers incorporated a variety of instructional strategies including computer-aided instruction. Student assessment was frequent and ongoing, and assessment data were used to make decisions regarding individual student programs. Riverview district leaders anticipated that students participating in Project Excel would demonstrate academic improvement. The
goal of this study was to examine the mathematics achievement of Title I-eligible students who participated in Project Excel. Three research questions were addressed during this investigation:

1. What is the significance, if any, of preexisting differences in mathematical abilities of the participating groups in this study?

2. What is the relationship between participation in Title I programs, specifically the Project Excel program, and mathematics achievement?

3. What is the relationship between participation in Title I programs, specifically the Project Excel program, and closing the achievement gaps between low-income and moderate/high-income students and non-Asian minority and White students?

Summary of Findings

Research question one. In the first investigation, the researcher determined the significance of preexisting differences in mathematical abilities prior to participation in Project Excel. All students took the DBA pretest during the first week of school prior to being invited to participate in the Project Excel program. To address this research question the researcher compared the DBA pretest scores of: (a) Project Excel participants to those of all non-participants, and Title I-eligible non-participants, (b) non-Asian minority participants to the DBA pretest scores of White non-participants, and (c) lower SES participants to the DBA pretest scores of moderate/higher SES non-participants.
The findings indicated that in each comparison, the participants scored significantly lower than did the non-participants on the DBA pretest, thereby confirming the existence of pre-participation differences in mathematical ability levels between the groups and the existence of SES and minority achievement gaps. The null hypothesis was rejected for all four cases.

Research question two

District leaders anticipated that participation in the Title I-funded Project Excel program would improve the academic performance of at-risk students, and they hoped this assumption would be supported by statistical evidence that the participants (Group A) performed better on the mathematics sections of the DBA posttest, TerraNova, and NJ ASK than did Title I-eligible non-participants (Group B). To address the second research question, the researcher compared the mathematics achievement of Project Excel participants to that of: (a) all non-participants, (b) Title I-eligible non-participants, and (3) district, state, and national norms.

The findings from the first comparison provided three main findings. First, Project Excel students scored significantly lower on the mathematics sections of the DBA posttest, TerraNova, and NJ ASK than did all non-participants. Second, although Project Excel participants scored significantly lower on the DBA posttest than did Title I-eligible non-participants, there was no statistically significant difference in performance between the two groups on the TerraNova and NJ ASK. These findings showed that Project Excel participants performed at the same level as did Title I-eligible non-participants on the TerraNova and the NJ ASK. Third, although Project Excel participants scored
significantly lower than the DBA posttest district norm and NJ ASK state norm, these students scored significantly higher than the TerraNova national norm.

The null hypotheses were rejected for the following comparisons: (a) all three investigations comparing the performance of the participants to that of all non-participants, (b) the investigation where Project Excel students performed significantly lower than did the Title I-eligible students on the DBA posttest, and (c) two investigations where the Project Excel participants scored significantly lower than the DBA posttest district and NJ ASK state norms. The null hypothesis was retained for the two cases where there was no statistical difference between the performance of Project Excel participants and that of eligible non-participants on the TerraNova and NJ ASK. The null hypothesis was rejected for the case where Project Excel participants performed significantly higher than the TerraNova national norm. To summarize the nine comparisons, six indicated that Project Excel participation was related to lower performance, two indicated that participation was related to no statistical differences between groups, and one indicated that participation was related to higher performance. Some evidence of improved academic performance was generated by three of the nine comparisons studied.

Although Project Excel participants scored significantly lower on the DBA pre- and posttests than did all non-participants and Title I-eligible non-participants, there was no significant difference in the mean pre-and posttest gain scores between the groups. This finding showed that Project Excel participants performed at the same rate as did other non-participants, thus showing that the participants experienced the same degree of growth as all other students over the course of the school year.
Research question three. Leaders anticipated the Project Excel program would help close the SES achievement gap between the economically disadvantaged students, most of whom were Hispanic and Black, and the middle/upper class students, most of whom were White. Although the administrative team did not anticipate that the economically disadvantaged students would overwhelmingly outperform the middle/upper class students, they did hope to obtain evidence to support the theory that participation in the extended-day program would reduce the gap between the two groups. To address the third research question, the researchers made two comparisons of mathematics achievement: (a) the performance of non-Asian minority participants to that of White non-participants and to district, state, and national norms and (b) the performance of lower SES participants to that of moderate/higher SES participants and to district, state, and national norms.

The findings from the third investigation showed mixed results. First, non-Asian minority and lower SES Project Excel participants scored significantly lower on the DBA posttest, TerraNova, and NJ ASK than did White and moderate/higher SES non-participants. Second, non-Asian minority and lower SES participants scored significantly lower than the DBA posttest district norm and the NJ ASK state norm than did White and moderate/higher SES non-participants. However, third, there were no statistically significant differences between the performance of non-Asian minority and lower SES participants and the TerraNova national norm. The non-Asian minority and lower SES participants performed equally as well on the TerraNova as did the average students across the nation.
The null hypothesis was rejected for the following comparisons: (a) all three cases where non-Asian minority participants performed significantly lower on the DBA posttest, TerraNova, and NJ ASK than did White non-participants, (b) all three cases where lower SES participants performed significantly lower on the DBA posttest, TerraNova, and NJ ASK than did moderate/higher SES non-participants, (c) two cases where non-Asian minority participants scored significantly lower than the DBA posttest district norm and NJ ASK state norm, and (d) two cases where lower SES participants scored significantly lower than the DBA posttest district norm and NJ ASK state norm. The null hypothesis is retained for the following comparisons: (a) one case where there was no significant difference in performance of non-Asian minority participants and the TerraNova national norms and (b) one case where there was no significant difference in performance of lower SES participants and the TerraNova national norms. Some evidence of closing the achievement gaps was generated by two of the 12 comparisons.

Although non-Asian minority and lower SES Project Excel participants scored significantly lower on the DBA pre- and posttests than did White and moderate/higher SES non-participants, there was no significant difference in the mean pre- and posttest gain scores between the groups. This finding showed that non-Asian minority and lower SES Project Excel participants performed at the same rate as did White and moderate/higher SES non-participants, and thus showed that there was no gap between the groups in terms of the same degree of growth achieved during the course of the school year.
Conclusions and Discussion

Do the overall results of this investigation suggest that participation in Project Excel was ineffective at improving mathematics achievement and closing the minority and SES achievement gaps? The answer to that question is a matter of perspective. Some leaders seek evidence of the effectiveness of Title I programs in the form of statistical proof that Title I students outperformed non-Title I students. Neekwosh, Cisneros and Sanchez as cited in Wong and Meyer (1998), and van der Klaauw (2005) found little statistical evidence in favor of Title I students performing significantly better than non-Title I students. In some cases a negative correlation was found which led researchers to conclude that participation in Title I programs was actually harmful for children.

In the present study the researcher compared the mathematics performance of Title I participants in relation to that of all non-participants. Findings from those comparisons add to the current body of research which indicated no significant improvement in performance when compared to that of non-participants, but rather a negative relationship: Project Excel participants performed significantly lower than did non-participants on the DBA posttest. TerraNova, and NJ ASK mathematics assessments.

The researcher cautions leaders in making policy decisions based on this and similar findings. There are drawbacks to evaluating Title I programs by comparing the performance of Title I participants with that of non-participants. This method of analysis is what the researcher deems as using the outperforming lens. Using such a lens as a means of judging a Title I program’s effectiveness may provide a cloudy view of what
may really be occurring. The outperforming lens fails to see the growth a population may
have made over a given period of time.

Consider the following analogy. Two cars are traveling along a highway. Car A, traveling at 65 mph is 50 miles ahead of Car B which is traveling at 45 mph. Suppose
Car B increases its speed to 60 mph and Car A continues to travel at 65 mph. The
distance between the two cars will still continue to increase over time. A researcher
using the outperforming lens to take a snap-shot picture of the performance of the two
cars one year later would find the gap between the two cars has widened and possibly
conclude that Car B has made no significant progress because it is still behind Car A.
This is a major flaw in using the outperforming lens to evaluate Title I effectiveness. In
order for the outperforming lens to yield a significant positive outcome as to the
effectiveness of Title I programs one of the following assumptions must hold true:

1. Non-Title I students must cease to improve while Title I students catch up;
2. Non-Title I students must experience a decrease in performance while Title I
   students catch up;
3. Title I students must experience performance growth at a faster rate of
   acceleration than do non-Title I students.

It seems unreasonable that the first two assumptions would be true. As for the third
assumption, researchers such as van der Klaauw (2005) proposed that Title I did not
improve achievement but rather prevented further widening of the various achievement
gaps. This prevention is what the researcher terms as sustaining gap equilibrium. Gap
equilibrium occurs when two comparable groups continue to perform at the same rate
thus maintaining the existence and size of the gap between the groups. Figure 19
provides a graphical representation of the D8A pre- and posttest performance of second- and third-grade students in this study. Note that the performance of the Project Excel participants parallels that of the non-participants, thereby sustaining gap equilibrium. Analysis of posttest gain scores showed that there was no statistical difference in the mean gain scores for Project Excel participants and non-participants, given the fact that both pre- and posttest means were significantly different between the two groups. Participants and non-participants were improving their performance at the same rate. This is a more holistic way than comparing pre- and posttest performance to examine the effectiveness of the program.

Few studies provided evidence that Title I has been effective at reducing the minority and SES achievement gap. Wong, Sunderman and Lee (1996) reported small gains at closing the SES achievement gap in Minneapolis schools, and Etheridge (2001) found small increments of gap shrinkage in Berkeley County, South Carolina schools. In reviewing the long-term benefits of participation in the Perry Preschool study, Schweinhart (2004) found “that adults at age 40 who had the preschool program had higher earnings, were more likely to hold a job, had committed fewer crimes, and were more likely to have graduated from high school than adults who did not have preschool” (p. 1).

Class size has been shown to close the SES achievement gap. Finn, Gerber, and Boyd-Zaharias (2005) noted that the graduation rates for low-income STAR students who participated in small classes for 3 or more years were similar to those of higher income students. Krueger and Whitmore (2001) found that Black STAR students who
Figure 19. Mean DBA mathematics pre- and posttest performance for participants and non-participants. Data collected from second- and third-grade students who participated in Project Excel (n = 35) and those who did not (n = 113). Note the parallel performance of the groups.
participated in small classes showed a significant increase in college entrance examination test-taking which resulted in a reduction of the Black/White achievement gap in college entrance exam test-taking by 54%. In the present study, non-Asian minority participants and lower SES participants exhibited mathematics performance similar to the TerraNova national norm. This is a small contribution to the research which suggested that Title I may assist in closing the minority and SES achievement gaps.

Does the lack of strong evidence that supports that minority and lower SES Project Excel participants are performing significantly better than are White and moderate/higher SES students suggest that the Project Excel program had no influence on closing the achievement gaps? To address the issue, the researcher revisits the theory of gap equilibrium and applies it to the minority and SES achievement gaps. There were no significant differences between the DBA pre- and posttest gain scores of non-Asian minority participants and those of White non-participants and between the gain scores of lower SES participants and those of moderate/higher SES non-participants. These at-risk students increased their performance from pre- to posttest at a similar rate as did White and moderate/higher SES students. Figures 20 and 21 show graphical representations of DBA pre- and posttest performance based on race and SES. Note the parallel performance between the groups, suggesting the gap between the groups has been maintained.

If Title I programs contribute to gap equilibrium, one must question whether the performance of the participants would continue to parallel that of non-participants in
Figure 20. Mean DBA mathematics pre- and posttest performance categorized by race.

Data collected from second- and third-grade non-Asian minority students who participated in Project Excel (n = 24) and White students who did not participate (n = 65). Note the parallel performance of the groups.
Figure 21. Mean DBA mathematics pre- and posttest performance categorized by SES. Data collected from second- and third-grade lower SES students who participated in Project Excel (n = 27) and moderate/higher SES students who did not participate (n = 84). Note the parallel performance of the groups.
absence of participation in the program. Cooper et al. as summarized in Borman (2003), Entwisle and Alexander as summarized in Edwards et al. (2001), and David and Pelavin (1977) as summarized in Borman and D'Agostino (1996) previously concluded that the small gains made by Title I students participating in special programs were lost during the summer when no program was provided. Therefore, one might speculate that in the absence of the extended-day program, the gaps between the minority and White students and the lower SES and moderate/higher SES students would widen over time. District leaders should take care in using the lack of statistical evidence that Title I students performed significantly better than non-participants as a basis for altering or canceling a Title I program. Such decisions to cancel a program that is preventing further widening of the achievement gaps pose potentially detrimental effects and serve to further propagate the existing gaps.

In chapter 4, the researcher discovered and reported several trends in the performance data. First, the DBA posttest was used to analyze performance comparisons in seven instances. In all seven comparisons, the Project Excel participants performed significantly lower on the posttest than did the comparable group. A similar pattern emerged from the NJ ASK performance data. In six of the seven cases, Project Excel participants performed significantly lower on the NJ ASK than did the comparable group. The TerraNova performance data yielded mixed findings. Project Excel participants performed significantly lower than did the comparable group in three of the seven comparisons, and there was no statistical difference in performance in three of the seven comparisons. Project Excel students scored significantly higher than the TerraNova
national norm. What accounts for these differences in performance on the multiple assessments?

One explanation may be that the TerraNova test has a higher reliability and lower standard error of measurement (SEM) than do the NJ ASK and DBA posttests. Tables 4 and 5, presented in Chapter 3, highlighted the reliability and SEM for the NJ ASK and TerraNova tests. Tienken and Wilson (2006) cautioned district leaders in using the NJ ASK as the sole tool for making decisions regarding programs. They stated:

Lower levels of reliability mean an increase in score error and a decrease in instrument precision. The lack of instrument precision can produce results that lead district administrators and teaching staff to draw faulty conclusions about the effectiveness of programs and curricula. A minimum reliability coefficient of at least .85 (Frisbie, 1988) should be used in making high-stakes decisions about students, although an argument can be made for a minimum of .90. (pp. 6-7)

Each reliability estimate for the four mathematics content areas tested on the NJ ASK 3 and 4 was less than 0.74 (see Tables 4 and 5).

Another explanation may be found by examining grade-level trends. Second- and fifth-grade students took the TerraNova test and an examination of the fifth-grade TerraNova data revealed that fifth-grade Project Excel participants had more positive gains than did second-, third-, and fourth-grade participants. The results from the fifth-grade comparison studies are summarized as follows:

1. Fifth-grade Project Excel participants performed equally as well on the TerraNova as did all non-participants and Title I-eligible non-participants (p > .05).
2. Fifth-grade lower SES participants performed equally as well on the TerraNova as did the moderate-upper-SES participants ($p > .05$).

3. Fifth-grade Project Excel students scored 16 points higher than the TerraNova national norm ($p = .006$).

4. Lower SES fifth-grade participants scored 15 points higher than the TerraNova national norm ($p = .014$).

5. Non-Asian minority fifth-grade participants scored 9 points higher than the TerraNova national norm ($p = .026$).

The small sample sizes of all Project Excel participating fifth-grade students ($n = 26$), fifth-grade non-Asian minority students ($n = 15$), and fifth-grade lower SES students ($n = 10$) precluded the researcher from presenting strong statistical evidence in the data analysis section; however, the observable trends provide evidence of positive academic achievement that district leaders should explore. What could the trends in the fifth-grade performance data be attributed to? Anecdotal notes taken from district records and qualitative reports provide some insight.

First, 100% of the fifth-grade Project Excel participants were instructed by Project Excel teachers who also taught fifth-grade math during the day. Perhaps fifth-grade Project Excel instructors who also taught fifth-grade math during the day had more knowledge about the essential math skills the students needed to be successful at that level than did Project Excel instructors who did not teach the same grade level as the students they instructed in Project Excel.

Second, 76% of the fifth-grade Project Excel participants, 0% of fourth-grade participants, 50% of third-grade participants, and 57% of second-grade participants were
instructed by the same teacher for daily math instruction and for Project Excel. These teachers had more knowledge of their students’ strengths and weaknesses as well as knowledge of what instructional strategies best worked with each student. In their review of Schwendiman and Fager's research, Dodd and Wise (2002) concluded that a primary reason for the success of after-school programs was because the regular classroom teachers were the primary instructors in the after-school program. These teachers were familiar both with the curriculum and the students and the difficulties they were experiencing.

The Project Excel instructors who taught their own students wrote more detailed descriptions in the student portfolios and noted more positive gains than did those teachers who did not teach their own students. Teachers who instructed their own students recorded more parental contact time than did teachers who did not instruct their own students. Further research is needed to determine the significance of this association.

Third, the extended-day program has been in existence for at least 5 years, although attendance records could only be produced for the last 3 years. The average years of participation on record was 2.6 years for fifth grade, 2.0 years for fourth grade, 1.5 years for third grade, and 1.2 years for second grade. This was the first year of participation in the program for most second-grade students and the third year or more for most fifth-grade students. In their research on class size, Finn et al. (2005) noted that low-income primary grade students who participated in smaller classes for 3 or more consecutive years or more experienced sustained and lasting benefits including higher high school graduation rates. Finn, Fox, McClelan, Achilles, and Boyd-Zaharias as cited
in Finn (2006) concluded that students who had sustained consecutive participation in
small class size in the primary grades were more likely to take more and higher levels of
foreign language and mathematics during high school. Krueger and Whitmore (2001)
provided evidence that small class size participation for 3 or more consecutive years had
a significant impact on improving college entrance exam test-taking experiences of Black
students and reduced the Black/White entrance exam achievement gap. The finding that
sustained participation is needed to produce statistically significant gains is relevant in
the present study. Findings from fifth-grade comparisons provide evidence in favor of
consecutive year Project Excel participation. District leaders should consider the benefits
of prolonged participation in Project Excel.

Fourth, the most significant difference in mathematics performance between fifth-
grade Project Excel students and that of eligible non-participants was found to be on the
computation and estimation section (classified under the number sense and numerical
operations standard of NJCCCS) of the TerraNova. Project Excel students scored over
10 points higher (scale of 100 points) on the computation section than did Title-I eligible
students ($p = .008$), and when their performance was compared to that of ineligible non-
participant, there was no statistically significant difference found. Was the strong
emphasis of an average of 79% of instructional time (fifth-grade teachers only) spent on
case topics related to this performance? District leaders should explore this avenue.

Fifth, what impact did computer-assisted instruction have on the findings?
Although the initial intent of the Project Excel program was to dedicate approximately
25% of the instructional time for students in Grades 3 through 5 to computer-assisted
instruction, most teachers noted problems using and/or accessing the computers and
opted to exclude them from the program. The fifth-grade teachers used the computers more frequently than did any other grade-level instructors. Many noted using the computers to reinforce basic skills and provide computational practice. In a meta-analysis of research conducted on computer-assisted instruction, Snow (2003) concluded that "the research supplies strong evidence that computer-assisted instruction is an effective strategy for meeting the needs of at-risk and low-performing students" (p. 51).

Did the enhanced time for computer-assisted instruction have an impact on the overall results? District leaders need to consider strongly the benefits of computer-aided instruction as improving the achievement of at-risk students.

Recommendations for Practice, Policy and Future Research

Longitudinal cohort analyses: In the present study, the researcher examined the mathematics achievement of Title I students who participated in the Project Excel program during the 2004 - 2005 school year. There were a total of 21 comparisons of the performance on the DIBA posttest, TerraNova, and NJ ASK mathematics assessments of Project Excel participants and various comparison groups. The results of the one-year investigation reflect the infancy of program implementation and evaluation. There was one statistically significant finding that showed that Project Excel students performed better than did Title I-eligible non-participants, two findings where Project Excel students performed the same as did eligible non-participants, and two findings where Project Excel students performed the same as state and national norms. The practice of Title I program evaluation should be an on-going longitudinal analysis of the performance of cohorts of children. The one-year snapshot study may not provide sufficient information
for district leaders to evaluate time-on-task, a component of the theoretical framework of
this study (Figure 1) and an opportunity-to-learn (OTL) variable. Berliner (1990), Cotton
(1989), Dodd and Wise (2002), Peterson (2005), and WestEd (1998) reported the benefits
of extended time-on-task, yet how does one quantify the effect of consecutive years of
participation in extended-day programs?

Much information regarding the lasting benefits and enduring effects of
participation in small class sizes for 3 or more consecutive years was learned from post-
STAR studies, and district leaders should consider the benefits of conducting similar
longitudinal studies to access the lasting benefits and enduring effects of participation in
extended-day programs. District leaders should continue to keep data on the
mathematical performance of new and current Project Excel participants to determine if
there are any lasting effects of participation over multiple years. Researcher van der
Klazuw (2005) said it well:

One potential reason for our failure to find a beneficial effect of Title I funding
may be that we are only estimating a short term effect, in the year immediately
after gaining eligibility. It may take time to make Title I-funded remedial
education programs operational and effective, and its implementation may in fact
initially have a disruptive effect on students. (p. 21)

District leaders may look at these results with skepticism because of the lack of
strong significant gains in mathematics achievement but the researcher cautions leaders
to invest more time in the program. MacLean (2003) noted:

The value of involvement in an intervention program can only truly be measured
over the long-term... From an administrator’s point of view, an early intervention
program may appear to be costly in terms of personnel units, time, and materials, but if the long-term success of children is observed as a result of involvement in such a program, then the initial cost is lessened by the long-term benefits of helping children be successful learners. (p. 173).

Further results from STAR showed that (a) early intervention and (b) duration (3-4 consecutive years) were required for (1) gains and (2) lasting benefits (Finn et al., 2001).

Program evaluation. "Given the goal of Title I is to increase student achievement, the most important indicators of implementation are in classrooms" (Taylor & Teddlie, 1999, p. 304). Program evaluation and implementation fidelity were two critical elements of the theoretical framework (Figure 1) for the present study, yet IRB approval prohibited the researcher from conducting frequent, thorough, and efficient classroom observations of Project Excel program implementation because of the potential for pressure from the supervisory nature of the researcher. Brophy as summarized in Muijs and Reynolds (2000), Parker and Gerber (2000), and Huang (2001) reported the benefits of quality instruction which included techniques such as linking new and prior knowledge, incorporating constructivist approaches, matching instructional styles to learning styles, and using inquiry-based strategies. This type of instruction is best evaluated when observed. District leaders should conduct frequent classroom walk-throughs for the purpose of insuring implementation fidelity. Walk-throughs provide opportunities to observe and evaluate other components outlined in the theoretical framework (Figure 1) for the present study and OTL variables such as instructional materials (modified curriculum) and instructional practices and methodologies.
One limitation of the present study was the absence of consistent data collected from teachers. All student achievement data were coded by a confidential secretary and given to the researcher during the summer following participation in the program. In following IRB recommendations, the researcher could not directly supervise the collection of data from teachers, and thus some instructors never administered the DBA posttest or never submitted their DBA posttest results to the confidential secretary. District leaders should incorporate an interactive system of program evaluation that allows for frequent and close monitoring of all aspects of the program including supervision of instruction and submission of data necessary to evaluate the program’s effectiveness at the building level.

IRB regulations precluded the researcher from gathering valuable data from interviews with Project Excel participants, teachers and parents regarding the effectiveness of the program at meeting its goals. Huang (2001) suggested, “By creating an evaluation plan that focuses on clear goals, program managers set a course for ongoing improvement in which goals and plans are continually monitored and improved over time” (p.57). The evaluation plan should not be an afterthought but rather “evaluation starts at the beginning of a program – even during the program planning” (Krempel & Achilles, 1979, p. 20). Krempel and Achilles explained that a thorough evaluation should include assessments of program initiation evaluation (planning), program operations (process) and program outcomes (goal attainment).

Early intervention. Rothstein (2004a, 2004b) noted that the minority and SES achievement gaps were formulated well before the child entered formal schooling. Phillips, Crouse and Ralph as summarized in Borzaan (2003) “found that about half of the
(B)lack-(W)hite achievement gap at the end of high school can be attributed to the fact that African American students begin school with fewer skills than (W)hite students" (p. 50). Recent litigation has supported the need for early intervention. The Abbott v. Burke ruling required the State of New Jersey to "implement a high quality preschool program for all three- and four-year-old children residing in the State’s highest poverty school districts" (Starting at 3, n.d.). Although Riverview is not one of the Abbott districts, the district does offer a voluntary half-day preschool program for all 3- and 4-year-olds. Experts Barnett of Rutgers University and Odden of the University of Wisconsin-Madison recommended full-day preschool programs (Starting at 3, n.d.).

A preschool program is a good start and extended time-on-task could provide the added benefit needed to close the "30 million word" gap (Hart & Risley, 2003). However, by the time pre-school at-risk students are eligible to take advantage of extended time by participating in Project Excel they have already had 3 to 4 years of public schooling and 7 to 8 years of differential experiences. District leaders should put forth a concerted effort to provide Title I-funded intervention programs in the early grades (P – 3) to combat these established differences. Running extended-day programs can be rather costly; however, the price of early intervention far outweighs the costs increased high school drop-out rates.

Mandatory participation. "Raising the achievement of lower-class children, narrowing the gap in cognitive achievement and non-cognitive skills between these children and those from the middle class, are more ambitious undertakings than policy makers today acknowledge" (Rothstein, 2004b, p. 149). School reform alone is not the solution; however, improving the ways educators address the achievement gaps is a start.
Project Excel is one such initiative that addressed the minority and SES achievement gaps, yet the program had some challenges that might be addressed by reviewing district policy.

One challenge faced by district leaders was parental choice. Parents had the option to choose if they wanted their child to participate in the Project Excel program. In many cases a choice not to participate was made due to constraints in providing transportation and childcare after the program dismissed. Perkins-Glough (2006) noted that schools that were identified by the Education Trust as being successful in improving the performance of struggling students had several characteristics. "They identified struggling students early and required them to get the help they needed" (p. 88, emphasis added). District leaders should consider a policy of mandatory or at least strong recommendations for enrollment in intervention programs. The goal of such a policy would be to serve all children who desperately need these services. District leaders should reach out to parents who opted not to have their child participate and urge them to consider both the district leaders' and the parents' responsibilities to help the child. Transportation and childcare considerations should be included as part of the planning and program implementation processes. Accountability measures must also be addressed to ensure that the services a child receives are meeting the child's needs. This requires a commitment of district resources to fund these endeavors and in most cases providing matching funds to supplement the Title I allocated resources.

Small class size. Education is a fluid system. A change in one area may have a rippling effect in another area; therefore, district leaders need not look at a single method of closing the achievement gap but rather look at the big picture and the system as a
whole. Research on extended-day programs shows that these programs, including the one in the present study, have shown small gains at improving student achievement of at-risk children. Students participating in the Project Excel program experienced an average class size of 3 to 4 students. Class size was a fundamental element of the theoretical framework (Figure 1) of the present study and an important OTL variable. The small class size allowed the instructors to focus more closely on individual student needs, yet Project Excel made up only a small fraction of the students' day, and these students often spent most of their day in classes of 20 or more students. The Tennessee STAR study provided strong evidence that consecutive years of participation in small class sizes in the primary grades are one of three highly relevant factors in improving the academic achievement of at-risk students (Achilles & Finn, 2001). The positive benefits from participation in STAR classes went well beyond the initial elementary grades, lasting into high school and beyond (Finn, 2006; Kuegler & Whitmire, 2001).

Additional studies were conducted that followed the STAR participants through Grades 4 through 7 (Lasting Benefits Study) and from eighth grade through high school and beyond (Enduring Effects Study) (Achilles, personal communication, January 7, 2006). Kuegler and Whitmire as summarized in Finn (2006) found that a "significant increase in test-taking [SAT and ACTs] was found for students who attended small classes...the benefit for Black students was substantially greater than for [White] students, reducing the [Black-White] gap in college-entrance - test-taking by 54%" (p. 1). Finn, Geber, and Boyd-Zaharias as summarized in Finn (2006) provided evidence that high school graduation rates of low-income children who participated in small classes for 3 or more years were as high as graduation rates of higher income children. Finn, Fox,
McClellan, Achilles and Boyd-Zaharias as summarized in Finn (2006) stated that participation in small classes for 3 or more years had a "significant positive impact on the amount of foreign language taken, and the highest levels taken in foreign language and mathematics" (p. 3).

One of the goals of the Project Excel program, to improve academic achievement has been proven to be met by implementing small class sizes in the primary grades. The researcher recommends that district leaders consider the combined approach of implementing small classes and providing extended-learning time for those who need it via Project Excel or a similar program. The benefits of such a combination will eliminate the need for expensive remedial programs and basic skills teachers later during the middle and high school grades. It is strongly recommended that district leaders implement mandatory small class size policies for the regular P - 3 educational program and extended-day programs.

Extended-year programs. Are the gains made by participation in extended-day programs lost over the summer? Borman and D'Agostino (1996), David and Pelavin as summarized in Borman and D'Agostino (1996), Borman (2003). Cooper et al. as summarized in Borman (2003), and Entwisle and Alexander as summarized in Edwards et al. (2001) have documented the deleterious effects of summer and vacation breaks on the academic progress of at-risk students. Findings from the present study resulted in few positive gains at improving academic performance and closing the achievement gap; yet these gains could easily be lost over the summer. “Low-income students have fewer out-of-school opportunities and resources to sustain the achievement gains that they make during school…as a result, over the summer, poor students tend to slip even farther
behind their more advantaged peers" (Borman, 2003, p. 51). These students need more opportunities to extend their learning through the summer months. Rothstein (2004b) stated:

We can't construct tests that separate learning during the school day from that in the afternoon or on weekends, but summer learning data are consistent with the achievement gap being entirely due to children's experiences before they enter kindergarten, in afternoons and on weekends, and during the summer. A strategy to close the achievement gap between lower-class and middle-class children cannot ignore these non-school hours. (p. 58)

District leaders should invest in extended-day programs as full-year initiatives and create policies that would provide extended-year schooling for all at-risk children. This may be a costly initiative, however, and investment in the early grades leads to some savings later if fewer middle and high school students required remedial services and attended summer school to repeat courses.

Future research. With increased regulations on how Title I money can be allocated, more districts are beginning to design their own extended-day supplemental programs while others are being forced to use out-of-district supplemental services. In either case district leaders want to be assured that the programs will meet the goals of increased student achievement and the NCLB requirement of 100% proficiency by 2014. Inherent in meeting that goal is effectively evaluating a program's success at meeting its goals. One of the challenges facing educators conducting research in schools is the difficulty to conduct truly randomized experiments.
In the present study, samples were chosen by teacher committees and further affected by parental choice to participate. An examination of parental consent forms revealed that 22 of the 32 parents (69%) of Title I-eligible students who scored at or above 57% on the DBA pretest opted not to have their children participate in the Project Excel program. The Project Excel group consisted of a larger representation of the lower-performing Title I students while the eligible non-participant group consisted of a larger representation of higher performing Title I students, thus illustrating the disparity between groups due to non-randomization and parent choice. To the extent possible, future researchers should randomly assign Title I-eligible students to participating and non-participating groups. This will insure a better balance within the groups.

In the present study, the researcher analyzed the data collected from two schools in one school district. The population of both schools was relatively small, and thus partitioning the sample led to small sample sizes. In some cases, as with the minority achievement gap it was difficult for the researcher to specifically examine the Black-White achievement gap, due to the small sample of Black students participating in the study. Future studies should incorporate more schools and/or districts, thereby increasing sample size. The demographic make-ups of the schools should be considered to insure ample representation of various minority and economically disadvantaged students.

An additional limitation of this investigation was the lack of observational data. All of the data regarding implementation fidelity was taken from the written notes the teachers provided in the student portfolio and district records. Although helpful information was obtained from these records, there was a disparity in the degree of detail written by the various instructors. The disparity led to questions that were pertinent to
the present study. For example, there was very little written on the implementation of PD strategies learned. Was there a connection between best practices and implemented practices? Frequent classroom observations provide data to answer that question. Frequent observations allow the researcher to triangulate data and check for bias in written teacher anecdotal notes.

Future research should include classroom observations as well as interviews and focus groups conducted with students, teachers, principals, and parents. The degree of parent involvement and the extent to which students spent their non-school hours is of particular interest, because these activities often influence learning. These observations, interviews, and focus groups provide rich data regarding implementation fidelity and other factors related to academic achievement.

The assessment instruments used in the present study posed an additional unexpected limitation. Although a pre- and posttest design was proposed, school-level and teacher decisions were made that precluded all teachers from administering the posttest. In the present study, there was no post-participation assessment given that was common to all students who participated in Project Excel, but rather students took the TerraNova in second and fifth grade or NJ ASK in third and fourth grade. The use of each test had potential drawbacks. TerraNova scores were reported as percentile rankings, a norm-referenced model. The norm-referenced model by design ranks students without considering the actual amount of content they know, but rather the relative amount of content they know compared to other students. This poses a potential problem in that a student scoring in the 35th percentile could have actually mastered 80% of the learning goals for that grade level; however, the ranking procedure does not reflect
that level of mastery. NIAST tests suffer from low reliability estimates, and high standard errors that deemed them inappropriate tools to use to truly assess what a child knows.

Future research should address the academic achievement issues using a battery of assessment tools that are criterion-referenced, aligned to the curricula taught, have high reliability and validity estimates and have low standard errors. These assessment tools should allow the researcher to compare the progress of a cohort of students from one year to the next.

Much of the funding provided under NCLB is tied to professional development (PD) and staff training, two OTL variables that were outlined in the theoretical framework (Figure 1) for the present study. Although research is inconclusive as to the effectiveness of PD at improving student achievement, Tierken and Achilles (2003) reported small positive gains associated with well-planned, small-group and ongoing PD. Little data on the nature and specificity of PD training Project Excel instructors engaged in were available from district records. Teachers working with these skills-fragile students should have well-planned, small-group, ongoing PD experiences that support effective instructional methodologies for improving student achievement. District leaders should include ongoing comprehensive PD specifically designed for teachers working with Title I students. A comprehensive PD program involves attending workshops, working with a co-teacher/trainer in the classroom, and providing time for reflection, evaluation, and plans for teacher growth and enhancement of skill level.

The outcome of this investigation has produced some additional questions to be addressed by future research:
1. What is the relationship between multiple-year participation in extended-day programs and student achievement? Most of the findings where there were positive gains and/or no significant difference between the participants and non-participants were found at the fifth-grade level. Many of these fifth graders have participated in the program for 3 or more years.

2. What optimal amount of time (1 hour, 2 hours…) spent in an extended-day program produces the greatest academic gains? Project Excel students received 60 to 75 minutes of additional instruction weekly. Would more time have led to more academic gains?

3. How important is the “teacher effect?” Most Project Excel students were taught by their regular math teachers. Did this have a positive, negative, or any impact on their performance?

**Summary**

There is a wealth of evidence that shows that various external factors contribute to the achievement gap, yet, according to Rothstein (2004b):

> policy makers almost universally conclude that these existing and persistent achievement gaps must be the results of wrongly designed school policies — either expectations that are too low, teachers who are insufficiently qualified, curricula that are badly designed, classes that are too large, school climates that are too undisciplined, leadership that is too unfocused or a combination of these. (p. 1)

The government has assigned 100% of the burden to school district leaders to wipeout the minority and SES achievement gaps by 2014. That is 8 years to completely eliminate
these gaps as measured by tests given in Grades 3 through 8 and 11. These critical assessments for determining a school’s fate are plagued with low reliability estimates and high standard errors of measurement. The clock is ticking and school districts face stiff penalties for failure to achieve the goal. Today’s district leaders must devise plans to meet the needs of the at-risk students with limited time and resources available. For over 40 years, America’s history has documented the struggle to close the achievement gaps, and although some progress has been made, much still remains.

Programs like Project Excel are a start at addressing the needs of Title I students. By providing early effective interventions and year-long opportunities to learn, educators can chip away at the established achievement gaps in hopes of eroding them. This is a huge job considering the small fraction of time that a person spends in school. “Each of us, if we live to be just 70 years old, spends only 5% of our lives in school. Considering the other 95% is spent ‘out there,’ the only really substantial thing education can do is help us to become continuous, lifelong learners” (Littky & Grabelle, 2004, p. 3).
References

Achilles, C. M. (2005, April). *Why hasn't class-size research been used appropriately (or even used)?* Paper presented to the American Educational Research Association conference in Montreal, Ontario, Canada.


De Kantor, A., Adair, J., Chang, A., & Stonehill, R. M. (2003). Ensuring quality and sustainability in after-school programs: How partnerships play a key role. In M. M. Brabeck, M. E. Walsh, & R. E. Latta (Eds.), Meeting at the Hyphen: Schools-
universities-communities-professions in collaboration for student achievement
and well being, 107th yearbook for the National Society for the Study of

reform. Chevy Chase, MD: Mid-Atlantic Consortium, Inc. Retrieved February


24-25.

Dodd-Collins, L. (2003). Framework for accountability of New Jersey Districts and
Schools. New Jersey Department of Education, Division of Student Services,
Office of Title I Program Planning and Accountability. Retrieved July 25, 2005,
from http://www.state.nj.us/njded/grants/ncb/policy/syp/account_files/frame.htm

intervention program. Journal of Education for Students Placed at Risk, 6(4),
411-426.

Etheridge, S. D. (2001). Title I schoolwide programs: District support for achieving

for publication.


28, 2006, from
http://www.highscope.org/Research/PerryProject/PerryAge40SumWeb.pdf


**Learning Goals: First Grade Everyday Mathematics**

**Secure Skills**
- Count up and back by 1s, starting with any number up to and including 20
- Count up to 20 objects
- Count by 2s to 20; count by 5s to 50

**Developing Secure Skills**
- Count up and back by 1s on the number grid
- Exchange pennies for nickels
- Tell time to the nearest hour
- Order and compare numbers to 22
- Tell time to the nearest half-hour
- Understand place value for 10s and 1s
- Know addition facts for +1, +0, doubles and sums of 10

**Developing Skills**
- Count by 5s to 40
- Count by 2s to 40
- Write numbers from 1 to 20
- Compare pairs of numbers less than 16
- Write and count tallies
- Calculate the values of various combinations of pennies and nickels
- Find complements of 10
- Solve simple addition and subtraction number stories
- Identify and complete patterns
- Solve simple addition and subtraction problems by skip counting on the number line and the number grid
- Identify numbers as odd and even
- Know the values of pennies, nickels, and dimes, and calculate the value of combinations of these coins
- Tell time to the nearest half-hour
- Solve simple number stories
- Find simple sums and missing addends
- Calculate the value of coin combinations
- Understand place value for longs and cubes
- Compare numbers using < and >
- Know +1, and +0, doubles and sums of 10 addition facts
- Find many names for a number
- Know addition facts
- Count sets of quarters, dimes, nickels, and pennies
• Identify and use patterns on a number grid
Beginning/Developing Skills

• Complete Frames-and-Arrows diagrams
• Use standard units for measuring length
• Learn simple addition facts
• Calculate the value of coin combinations—"P," "N," "D," and "Q"
• Sort and identify objects by attributes
• Identify polygons and know their characteristics
• Identify fractional parts of regions and sets with a focus on unit fractions
• Identify fractional parts of a region
• Learning Goal By Assessment Rubric Indicators

Beginning Skills

• Find missing numbers and/or the missing rule in "What's My Rule?" problems
• Measure objects to the nearest centimeters
• Understand digital notation for time
• Identify 3-dimensional shapes and know their characteristics
• Identify symmetrical figures
• Make change for amounts less than $1
• Solve 2-digit addition and subtraction problems
• Compare fractions less than 1
• Find equivalent fractions
Learning Goals: Second-grade *Everyday Mathematics*

**Secure Skills**
- Count by 2s, 5s and 10s
- Make tallies and give the total
- Know "easier" addition facts
- Construct fact families for addition and subtraction
- Complete simple Frames-and-Arrows diagrams
- Solve simple addition number stories
- Find equivalent names for numbers
- Identify place value in 2-digit and 3-digit numbers
- Show "P," "N," "D," and "Q" for a given amount
- Know all addition facts
- Know easy subtraction facts
- Add and subtract multiples of 10
- Identify 2-dimensional shapes
- Add three 1-digit numbers mentally
- Know complements of 10
- Count by 2s, 5s, and 10s and describe the patterns
- Find missing addends for the next multiple of 10
- Solve number-grid puzzles
- Plot data on a bar graph
- Shade a specified fractional part of a region
- Give the fraction name for the shaded part of a region
- Use a ruler, tape measure, and meter/yardstick correctly
- Read and write money amounts in decimal notation
- Use equivelated coins to show money amounts in different ways
- Use a calculator to compute money amounts
- Know exchange values of U.S. coins
- Know and express automatically the values of digits in 2-, 3- and 4-digit numbers
- Multiply numbers with 0 or 1 as a factor
- Tell time to 5-minute intervals
- Demonstrate calendar concepts and skills
- Compare quantities from a bar graph

**Developing/Secure Skills**
- Know "harder" addition facts
- Know "easier" subtraction facts
- Complete "What's My Rule?" tables
- Solve simple subtraction number stories
- Tell time to 5-minute intervals
- Draw line segments
- Add and subtract with multiples of 10
- Solve addition and subtraction number stories
• Measure to the nearest inch
• Measure to the nearest centimeter
• Identify equivalencies for inches, feet, and yards
• Construct multiplication/division fact families
• Multiply numbers with 2, 5, and 10 as a factor
• Learning Goal By Assessment Rubric Indicators
Developing Skills

- Find values of coin and bill combinations
- Know "easy" addition facts
- Identify place value for ones, tens, and hundreds
- Complete number sequences; identify and use number patterns to solve problems
- Find equivalent names for numbers
- Compare numbers: write symbols <, >, and =
- Know "harder" subtraction facts
- Solve Francis-and-Arrows problems having two rules
- Make change
- Know more difficult subtraction facts
- Devise and use strategies for finding sums of 2-digit numbers
- Devise and use strategies for finding differences of 2-digit numbers
- Estimate approximate costs and sums
- Read "F" on a thermometer
- Identify 3-dimensional shapes, such as rectangular prisms, cylinders, pyramids, cones, and spheres
- Identify symmetrical figures
- Find common attributes of shapes
- Identify parallel and nonparallel line segments
- Use the trade-first method to solve 2-digit subtraction problems
- Make ballpark estimates of exact answers
- Model multiplication problems with arrays
- Add three 2-digit numbers mentally
- Find missing addends for any multiple of 10
- Find the median (middle value) of a data set
- Understand fractions as names for equal parts of a region or set
- Understand that the amount represented by a fraction depends on the size of the whole (ONS)
- Shade a specified fractional part of a collection
- Give the fraction name for the shaded part of a collection
- Recognize equivalent fraction names
- Measure to the nearest 1/2 inch
- Measure to the nearest 1/2 cm
- Use appropriate units for measurement and recognize sensible measurements
- Find area concretely
- Find perimeter concretely
- Solve money stories involving change
- Estimate totals for "ballpark" check of exact answers
- Know and express automatically the values of digits in 5-digit numerals
- Estimate and solve addition and subtraction number stories with dollars and cents
- Solve 1-digit multiplication stories (multiples of equal groups)
- Solve simple division stories (equal sharing and equal groupings)
- Multiply numbers with 2, 5, or 10 as a factor
• Construct multiplication/division fact families
• Make difference and ratio comparisons
• Determines the median, maximum, minimum, and range of a data set

**Beginning/Developing Skills**
• Solve stories about multiples of equal groups
• Solve equal-grouping and equal-sharing division problems
• Identify equivalencies for mm, cm, dm and m
• Compare fractions

**Beginning Skills**
• Use parentheses in number models
• Use alternate names for times of day
• Know "harder" multiplication facts
• Determine the mode of a data set
Learning Goals: Third-grade *Everyday Mathematics*

**Secure Skills**
- Know basic addition facts
- Know basic addition and subtraction facts
- Complete fact and number families
- Solve addition and subtraction multidigit number stories
- Add multidigit numbers
- Subtract multidigit numbers
- Measure line segments to the nearest 1/4 inch
- Measure line segments to the nearest centimeter
- Know multiplication facts having 0 or 1 as a factor
- Read, write, and compare whole numbers up to 5 digits
- Identify place value in whole numbers up to 5 digits
- Know multiplication facts from the first set of Fact Triangles
- Identify right angles
- Identify and name 2-D and 3-D shapes
- Identify symmetric figures and draw lines of symmetry

**Developing/Secure Skills**
- Count by 10s and 100s
- Apply place-value concepts in 4-digit numbers
- Tell and show times to the nearest minute
- Count combinations of bills and coins and write the total in dollars-and-cents notation
- Find equivalent names for numbers
- Use basic facts to solve fact extensions
- Complete “What’s My Rule?” tables
- Know multiplication facts having 2, 5, or 10 as a factor
- Complete multiplication/division fact families
- Know multiplication facts from the first set of Fact Triangles
- Make a bar graph

**Developing Skills**
- Identify and use number patterns to solve problems
- Estimate answers to multidigit addition and subtraction problems
- Find the perimeter of a polygon
- Find the area of a rectangular region divided into square units
- Solve number stories involving equal groups by using multiplication
- Solve number stories involving equal sharing and equal grouping
- Know multiplication facts having 3 or 4 as one factor and 2 through 7 as the other factor
- Compare and order decimals
- Identify place value in decimals
- Read and write 1-and 2-digit decimals
• Identify, draw, and name line segments, lines, and rays
• Draw parallel and intersecting line segments, lines, and rays
• Draw angles at records of rotations
• Recognize and know square products
• Know multiplication facts from second set of Fact Triangles
• Solve extended multiplication facts to tens x tens
• Identify fractional parts of a set
• Identify fractional parts of a region
• Solve extended multiplication facts to hundreds x hundreds
• Solve number stories involving equal shares and equal groups
• Find the median of a data set
• Measure in centimeters and inches
• Know units of measure
• Make a frequency table
• Know multiplication facts

Beginning/Developing Skills
• Understand fraction and placement of parentheses in number sentences
• Make ballpark estimates for sums and products
• Identify fractions on a number line
• Find equivalent fractions
• Solve fraction number stories
• Use the partial-products algorithm or the lattice method to multiply multidigit numbers by 1-digit or 2-digit numbers
• Find factors of a number
• Interpret remainders in division problems
• Find the volume of rectangular prisms
• Find the mean of a data set
• Understands and uses the language of probability
• Uses fractions to record probability of events
• Uses random draws to predict outcomes
• Collects and organizes data for use in predicting outcomes
• Understands area model of probability and solves simple spinner problems

Beginning Skills
• Read, write, and compare 6- and 7-digit whole numbers
• Read and write 3-digit decimals
• Compare and order fractions
• Convert between mixed numbers and fractions
• Solve number stories involving positive and negative numbers
Learning Goals: Fourth-grade Everyday Mathematics

Secure Skills
• Name, draw, and label line segments, lines and rays
• Name, draw, and label angles, triangles, and quadrangles
• Identify and describe right angles and parallel lines and line segments
• Solve addition and subtraction facts
• Use the statistical landmarks maximum and minimum
• Have a successful strategy for subtracting multidigit numbers
• Have a successful strategy for adding multidigit numbers
• Read and write numerals to hundred-million; give the value of the digits in numerals to hundred-millions
• Give equivalent names for numbers
• Solve basic multiplication facts
• Understand the relationship between multiplication and division
• Draw and measure line segments to the nearest centimeter
• Use dollars-and-cents notations
• Compare large numbers
• Estimate sums
• Identify the whole for fractions
• Identify fractional parts of a collection of objects
• Identify fractional parts of regions
• Give equivalencies between hundredths-fractions, decimals and percents
• Use a calculator to rename any fractions as a decimal or percent
• Use a transparent mirror to draw the reflection of a figure
• Identify lines of symmetry, lines of reflection, reflected figures, and figures with line symmetry
• Solve rate problems, using rate tables as necessary

Developing Skills
• Identify properties of polygons
• Classify quadrangles according to side and angle properties
• Display data with a line plot, bar graph, or tally chart
• Use the statistical landmarks: median, mode, and range
• Solve open sentences
• Insert parentheses to make true number sentences; solve problems with parentheses
• Determine whether number sentences are true or false
• Use and explain strategies for solving addition and subtraction number stories
• Use a map scale to estimate distances
• Solve basic division facts
• Express metric measures with decimals
• Convert between metric measures
• Read and write decimals to thousandths
• Compare and order decimals
• Draw and measure line segments to the nearest millimeter
• Use personal references to estimate lengths in metric units
• Solve 1- and 2-place decimal addition and subtraction problems and number stories
• Solve extended multiplication facts
• Make magnitude estimates for products of multidigit numbers
• Solve multidigit multiplication problems
• Round whole numbers to a given place
• Read and write numbers to billions; name the values of digits in numerals to billions
• Have a successful strategy for solving whole-number division problems
• Express the remainder of a whole-number division problem as a fraction and the answer as a mixed number
• Interpret the remainder as division problems
• Name and locate points specified by ordered number pairs on a coordinate grid
• Learning Goal By Assessment Rubric Indicators
• Identify acute, right, obtuse, straight, and reflex angles
• Make turns and fractions of turns; relate turns and angles
• Use a circular protractor and a half-circle protractor to measure and draw angles
• Use and explain strategies for solving multiplication and division number stories
• Rename fractions with denominators of 10 and 100 as decimals
• Apply basic vocabulary and concepts associated with chance events
• Compare and order fractions
• Find fractions equivalent to a given fraction
• Use formulas to find area of rectangles, parallelograms, and triangles
• Find the perimeter of a polygon
• Estimate the area of a figure by counting unit squares and fractions of unit squares inside the figure
• Find a percent or a fraction of a number
• Give equivalencies between "easy" fractions (fourths, fifths, and tenths), decimals, and percents
• Translate figures
• Add positive and negative integers
• Estimate the weight of objects in ounces or grams; weigh objects in ounces or grams
• Solve cube-stacking volume problems
• Describe properties of geometric solids
• Find unit rates
• Calculate unit prices to determine which product is the "better buy"
• Evaluate the reasonableness of rate data
• Collect and compare rate data

Beginning Skills
• Use a compass and straightedge to construct geometric figures
• Use exponential notation to represent powers of 10
• Identify locations on Earth for which latitude and longitude are given; find latitude and longitude for given locations
• Add and subtract fractions
• Make and interpret scale drawings
• Use an estimation strategy to divide decimals by whole numbers
• Use an estimation strategy to multiply decimals by whole numbers
• Add integers
• Rotate figures
• Use a formula to calculate the volume of rectangular prisms
• Subtract positive and negative integers
• Find unit rates
Learning Goals: Fifth-grade Everyday Mathematics

Secure Skills
- Draw arrays to model multiplication
- Know basic multiplication facts
- Identify even and odd numbers
- List the factors of a number
- Find the sum and difference of multidigit whole numbers and decimals
- Identify the maximum, minimum, mode, and mean for a data set
- Identify place value in numbers to billions
- Know properties of polygons
- Define and create tessellations
- Know place value to hundredths
- Find and use data landmarks
- Convert among fractions, decimals and percents
- Convert between fractions and mixed or whole numbers
- Find common denominators
- Understand the concept of area of a figure
- Use a formula to find the area of rectangles
- Use formulas to find the area of polygons and circles
- Know the properties of geometric solids
- Find and identify factors of numbers
- Find the prime factorizations of numbers

Developing/Secure Skills
- Use a divisibility test to determine if a number is divisible by another number
- Identify prime and composite numbers
- Understand how square numbers and their square roots are related
- Make magnitude estimates
- Find the product of multidigit whole numbers and decimals
- Know place value to billions
- Estimate the measure of an angle
- Measure an angle to within 2°
- Identify types of angles
- Identify types of triangles
- Convert between fractions and mixed numbers
- Find equivalent fractions
- Convert among fractions, decimals, and percents
- Understand and apply exponential notation
- Identify number sentences; tell whether a number sentence is true or false
- Understand and apply the use of parentheses in number sentences
- Order and compare positive and negative numbers
- Use an algorithm to add mixed numbers
- Order and compare fractions
• Plot ordered pairs on a one-quadrant coordinate grid
• Identify the base and height of triangles and parallelograms
• Use a formula to find the area of triangles and parallelograms
• Solve one-step pan-balance problems
• Interpret mystery line plots and graphs
• Use formulas to find the volume of prisms and cylinders
• Solve ratio and rate number stories

**Developing Skills**
• Round numbers to designated places
• Determine angle measures based on relationships between angles
• Find the quotient and remainder of a whole number divided by a 1-digit whole number
• Find the quotient and remainder of a whole number divided by a 2-digit whole number
• Make magnitude estimates for quotients of whole and decimal numbers divided by whole numbers
• Interpret the remainder in division number stories
• Determine the value of a variable; use this value to complete a number sentence
• Order and compare fractions
• Convert between fractions and percents
• Draw a circle graph for a set of data
• Measure pieces of a circle graph; interpret a circle graph
• Add and subtract fractions with common denominators
• Add and subtract fractions with unlike denominators
• Understand how sample size affects results
• Find a common denominator
• Understand and apply powers of 10
• Understand and apply order of operations to evaluate expressions and solve number sentences
• Add and subtract positive and negative numbers
• Use an algorithm to multiply fractions
• Use an algorithm to subtract mixed numbers with like denominators
• Find a percent of a number
• Plot ordered pairs on a four-quadrant coordinate grid
• Understand the concept of volume of a figure
• Use a formula to find the volume of prisms
• Write algebraic expressions to describe situations
• Represent rate problems as formulas, graphs, and tables
• Use formulas to find circumference and area of a circle
• Distinguish between circumference and area of circle problems
• Find the greatest common factor of two numbers
• Find the least common multiple of two numbers
**Beginning/Developing Skills**
- Rename numbers written in exponential notation
- Add fractions with like denominators
- Construct stem-and-leaf plots
- Read and interpret stem-and-leaf plots
- Understand and apply scientific notation
- Use the Multiplication Counting Principle to find the total number of possible outcomes of a sequence of choices

**Beginning Skills**
- Find the prime factorization of numbers
- Write and solve open sentences for number stories
- Divide decimals by whole numbers with no remainders
- Write and solve number sentences with variables for division number stories
- Use an algorithm to multiply mixed numbers
- Solve two-step pan-balance problems
- Understand the relationship between the volume of pyramids and prisms, and the volume of cones and cylinders
- Find the surface area of prisms
- Understand how to find the surface area of cylinders
- Understand the concept of capacity and how to calculate it
- Use tree diagrams to find all possible ways a sequence of choices can be made
- Compute the probability of outcomes when choices are equally likely
Appendix B
Standard 1 – Number and Numerical Operations
• Number Sense
• Numerical Operations
• Estimation

Standard 2 – Geometry & Measurement
• Geometric Properties
• Transforming Shapes
• Coordinates of Geometry
• Units of Measurement
• Measuring Geometric Objects

Standard 3 – Patterns and Algebra
• Patterns and Relationships
• Functions
• Modeling
• Procedures

Standard 4 – Data Analysis, Probability, and Discrete Mathematics
• Data Analysis (Statistics)
• Probability
• Discrete Mathematics – systematic listing and counting
• Discrete Mathematics – vertex-edge graphs and algorithms

Standard 5 – Mathematical Processes
• Problem Solving
• Communication
• Connections
• Reasoning
• Representations
• Technology
Appendix C
The NJ ASK 3 contains 30 test items: 27 multiple choice items, six of which are non-calculator multiple choice items and 3 open-ended response questions. The total raw score points a student could earn is 33. The minimum number of raw points necessary to achieve "proficient" is 17/33. The minimum number of raw points necessary to achieve "advanced proficient" is 27 5/33.

- 30.3% (10) of the points assess Number and Numerical Operations
- 24.2% (8) of the points assess Geometry and Measurement
- 24.2% (8) of the points assess Patterns and Algebra
- 21.2% (7) of the points assess Data Analysis, Probability, and Discrete Mathematics

The NJ ASK 4 contains 37 test items: 32 multiple choice items, eight of which are non-calculator multiple choice items and 5 open-ended response questions. The total raw score points a student could earn is 43. The minimum number of raw points necessary to achieve "proficient" is 19.5/43. The minimum number of raw points necessary to achieve "advanced proficient" is 32.5/43.

- 30.2% (13) of the points assess Number and Numerical Operations
- 23.3% (10) of the points assess Geometry and Measurement
- 23.3% (10) of the points assess Patterns and Algebra
- 23.2% (10) of the points assess Data Analysis, Probability, and Discrete Mathematics