The Academic Impact Of Service-Learning On New Jersey Public High Schools

Andrea Lynne Dinan
Seton Hall University

Follow this and additional works at: https://scholarship.shu.edu/dissertations
Part of the Secondary Education and Teaching Commons

Recommended Citation
https://scholarship.shu.edu/dissertations/1489
THE ACADEMIC IMPACT OF SERVICE-LEARNING ON NEW JERSEY PUBLIC HIGH SCHOOLS
BY
ANDREA LYNNE DINAN

Dissertation Committee
Elaine Walker, Ph.D., Mentor
Daniel Gutmore, Ph.D.
Jeffrey Graber, Ed.D.
Ron Horowitz, Ed.D.

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Education
Seton Hall University
2005
ABSTRACT

THE ACADEMIC IMPACT OF SERVICE-LEARNING ON NEW JERSEY PUBLIC HIGH SCHOOLS

ANDREA L. DINAN

This quasi-experimental study examined the impact of service-learning on academic achievement of New Jersey public high school students. Average standardized test scores were gathered for the High School Proficiency Achievement (HSPA) and the SAT test. Statistics addressing student mobility and student participation in SAT testing were also gathered. A survey disseminated by the State Department of Education was utilized to identify schools with service-learning programming and those without. The sample of service-learning schools included both schools that indicated that more than 50% of the student population was involved in service-learning programs and schools that were nominated to take part in the National Service-Learning Leader Schools program. These schools were matched to more traditional schools that do not employ service-learning programming. Analysis of Covariance tests demonstrated significant variance in Community Service-Learning Schools mean test scores with the Community Service-Learning Schools mean scores found to be higher than those of Traditional Schools. Significant interactions between SES groupings and the CSL School variable were also noted in some cases. The highest socioeconomic grouping revealed the highest means and the lowest socioeconomic grouping labeled revealed the highest mean differences between CSL and traditional schools. In both cases, the middle socioeconomic grouping showed a small but higher mean score for Traditional Schools when compared to schools with CSL programming. When the same statistical tests were performed on mean test
scores of New Jersey Service-Learning Leader Schools (NJ Leader) and Non-Leader Schools, a statistically significant difference was found in the verbal and math portion of the HSPA, the verbal and math portion of the SAT, and total SAT scores. In all instances the NJ Leader Schools outscored the Non-Leader Schools. Nine of the ten Null Hypotheses were rejected as they relate to the three subsidiary questions.
ACKNOWLEDGMENTS

I would like to acknowledge my mentor, Dr. Elaine Walker, for her endless support and insightful advice both in the writing process and in the statistical analysis. Without her, this thesis would still be in draft stages.

I would also like to thank Dr. Daniel Gutmore, for his quick replies and quiet encouragement in this long process.

I would also like to show my appreciation for Dr. Jeifery Graber for continuing the tradition of encouraging and compassionate assistance in the thesis process.

I would like to express thanks as well to Dr. Ronald Horowitz, the original Learning in the Community Director who provided the motivation for my career goals.

Thanks also to the fabulous Cohort VII: 30 professionals always available for a quick laugh, cry, complaint or notes.

I would also like to thank Linda Rivera, Program Director for Learn & Serve America of New Jersey Department of State, for all of her assistance, boundless good humor, and strength in leading service-learning programming in the State of New Jersey.
DEDICATION

I would like to dedicate this work to all of the mentors that I have been fortunate enough to have helped me along in life. From the early years of grade school, Rebecca Limons took the extra time and energy to encourage my writing and my confidence. Throughout high school, my neighbor and boss Donald McAghon provided the moral support needed to continue to college. In college, I was fortunate enough to find Dr. Gary Olson who guided me towards graduate school and scholarly studies. Following graduate school, a providential meeting with Wayne Miesel, brought me to Princeton and inspired and ethic of service in me that has driven my career. Since coming to Princeton, I have been fortunate to meet John Thurber, a professional and personal inspiration who encouraged me to seek out my dream of a doctoral degree. I would also like to include my best friend and inspiration, José Arango who helped me to keep my eyes on the goal of completing this program. Lastly, most of all, I would like to dedicate this work to my mother, Sharon Brynilson who has always been there for me and has pushed me to reach out for my dreams.
TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................... iii
DEDICATION ............................................................................................................................ iv
LIST OF TABLES ................................................................................................................... vii
LIST OF FIGURES .................................................................................................................. x

CHAPTER I: INTRODUCTION ................................................................................................. 1
  Conceptual Framework ...................................................................................................... 3
  Statement of the Problem ................................................................................................... 4
  Purpose of the Study .......................................................................................................... 8
  Significance of the Problem ............................................................................................... 9
  Research Questions ........................................................................................................... 10
  Hypotheses ......................................................................................................................... 10
    Subsidiary Question One .................................................................................................. 10
    Subsidiary Question Two ............................................................................................... 11
    Subsidiary Question Three ........................................................................................... 12
  Definition of Terms .......................................................................................................... 15
  Limitations .......................................................................................................................... 16
  Summary ............................................................................................................................. 17
  Organization of the Study ................................................................................................. 18

CHAPTER II: LITERATURE REVIEW .................................................................................... 19
  Historical Overview .......................................................................................................... 19
  Philosophy of Service-Learning ....................................................................................... 24
  Cognitive Thinking and Service-Learning ....................................................................... 27
  Experiential Education and Service-Learning ................................................................. 36
    Direct Experience .......................................................................................................... 36
    Reflection ......................................................................................................................... 37
  Service-Learning Research .............................................................................................. 40
    College-Based Research ............................................................................................... 40
    High-School Based Research ......................................................................................... 44
  Summary ............................................................................................................................. 50

CHAPTER III: METHODOLOGY ............................................................................................ 52
  Research Design ................................................................................................................ 52
  Problem Statement .......................................................................................................... 56
  Research Questions .......................................................................................................... 56
  Population .......................................................................................................................... 57
  Data Collection ................................................................................................................ 61
  Data Analysis ..................................................................................................................... 62
LIST OF TABLES

Tables

1. New Jersey Leader Schools (NJL) and their DFG Affiliation ........................................... 59
2. New Jersey Community Service-Learning High Schools (CCL) ............................................ 69
3. Distribution of Redefined DFG Designations in Re-Coding for Study ........................................ 61
4. Distribution of Schools by DFG ..................................................................................... 61
5. Data Sources and Analysis Chart .................................................................................... 64
6. Descriptive Statistics for Schools ..................................................................................... 68
7. Independent t-Test Statistics for Schools ............................................................................ 69
8. Descriptive Statistics for Schools and Demographic Variables ............................................ 70
9. Independent t-Test Statistics for School Demographic Variables ........................................... 76
10. Test Characteristics of the SAT .......................................................................................... 71
11. Distribution of Redefined DFG Designations in Re-Coding ................................................. 77
12. Descriptive Statistics of Test Scores and State Mean Test Scores ......................................... 79
13. Group HSPA Verbal Scores ............................................................................................... 82
14. Analysis of Covariance for HSPA Verbal Scores ................................................................ 83
15. Group HSPA Math Scores ................................................................................................. 84
16. Analysis of Covariance for HSPA Math Scores ................................................................ 85
17. Group SAT Verbal Scores ................................................................................................. 86
18. Analysis of Covariance for SAT Verbal Scores ................................................................ 86
19. Estimated Marginal Means Statistics for SAT Verbal Test ................................................. 88
20. Estimated Marginal Mean Statistics by School for SAT Verbal Test ...................................... 88
21. Estimated Marginal Means for ReDFG by School on SAT Verbal Test...89
22. Group SAT Math Scores .................................................................90
23. Analysis of Covariance for SAT Math Scores ..................................92
26. Estimated Marginal Means for ReDFG by School on SAT Math Test ....95
27. Group SAT Total Scores .....................................................................97
28. Analysis of Covariance for SAT Total Scores ....................................99
29. Estimated Marginal Means Statistics for SAT Total ..........................100
30. Estimated Marginal Means Statistics by School for SAT Total ..........101
31. Estimated Marginal Means for ReDFG by School for SAT Total ........101
32. Group HSPA Verbal Scores by Leader Schools (LS) .........................105
33. Analysis of Covariance for HSPA Verbal Scores by LS ....................105
34. Group HSPA Math Scores by Leader Schools (LS) ............................107
35. Analysis of Covariance for HSPA Math Scores by LS ......................107
36. Group SAT Verbal Scores by Leader Schools (LS) ............................109
37. Analysis of Covariance for SAT Verbal Scores by LS ......................109
38. Estimated Marginal Means Statistics for SAT Verbal Test by LS ........111
39. Estimated Marginal Means by ReDFG for SAT Verbal Test by LS ......111
40. Group SAT Math Scores by Leader Schools (LS) ..............................112
41. Analysis of Covariance for SAT Math Scores by LS .......................113
42. Estimated Marginal Means Statistics for SAT Math Test by LS ........114
43. Estimated Marginal Means by ReDFG on SAT Math Test by LS ..........115
44. Group SAT Total Scores by Leader Schools (LS) .........................116
45. Analysis of Covariance for SAT Total Scores by LS .....................116
46. Estimated Marginal Means Statistics for SAT Total by LS ..............117
47. Estimated Marginal Means Statistics by ReDFG for SAT Total by LS..118
48. Summary of Findings .............................................................120
LIST OF FIGURES

Figures

1. Group Verbal SAT Scores ................................................................. 90
2. Group Math SAT Scores ............................................................... 96
3. Group Total SAT Scores ............................................................... 102
CHAPTER I

INTRODUCTION

Today's social forces have dramatically challenged the traditional teaching and learning that occurred within the school building and was quite common twenty years ago. The opportunities for students are boundless, and to serve students' civic and academic development, schools must possess the programs that facilitate connections outside of the traditional classroom. Success in citizenship and academic learning has been demonstrated in many service-learning programs. It is often these activities that produce our most valuable citizens and community leaders. The preparation of students to be productive while fostering values is necessary in order to create active and compassionate members of society.

W. K. Kellogg Foundation and the Ewing Marion Kauffman Foundation presented results from a Roper Starch Survey (2000) that focused on determining the public vision of K-12 education in the United States. A sample that included 1,013 adults was collected via telephone using Random Digit Dialing. Findings showed 49% of Americans chose education as the highest priority for the United States, making education by far the priority on which there is the greatest public consensus. Other results included that 83% strongly agreed, "A good education is much more than just learning to read, write and do math," and 94% agreed that people not having the education or skills they need to succeed is a serious problem facing the United States.
These findings begin to illustrate the changing perceptions of education in the United States. Today, reading, writing, and mathematics instruction is not enough. The public opinion of education is that students need to be prepared for life after school in a democratic society.

The Kellogg and Kauffman Survey also included questions regarding service-learning and its purpose in the classroom. From the adults surveyed 90% believed service-learning will help students build skills they need to be successful later in life and 89% believed service-learning will encourage active citizenship and community involvement among students.

These results suggest that service-learning is viewed as being able to encourage positive effects. It is worth our time to evaluate it closely. While it may not be a palliative for all that is wrong in schools, service-learning can become a valued pedagogy and philosophy.

Service-learning is an educational method that connects the curriculum with acts of service performed by the students. The underlying goals are to strengthen the learning practice with real world activities in an effort to create, at the same time as stimulating, a sense of citizenship within the student while serving the community. Service-learning differs from community service in its core themes. While community service is often focused on a social issue and is commonly fostered in clubs and activities, service-learning is part of teaching pedagogy and intrinsically related to academic disciplines.

A standard definition of service-learning is as follows (Learning In Deed, 2003):

Service-learning combines service objectives with learning objectives with the intent that the activity changes both the recipient and the provider of the service.
This is accomplished by combining service tasks with structured opportunities that link the task to self-reflection, self-discovery, and the acquisition and comprehension of values, skills, and knowledge content.

The service-learning curriculum attempts to address real world issues in combination with an ethic of service. Studies continue to prove that this philosophy assists students with understanding and learning about their community and themselves (Follman, 1999; Melchior, 1999; Morgan, 2000; Morgan & Streb, 1999; Shumer, 1994).

Conceptual Framework

This study was guided conceptually by the writing of Dewey. Historically service-learning was first discussed largely in the works of John Dewey. Dewey (1916) believed that learning was facilitated by experience, and it would be largely through experiential education rather than traditional methods that students would learn the skills to participate in American democracy. Dewey also believed that activity focusing on the welfare of others served to inspire both academic and social development in students (Dewey 1916, 1933). The legacy of Dewey includes the idea that authentic learning occurs best when focusing on genuine problems in the community. Preparing students for active participation in the democratic system includes the solution of core community problems. Harkavy (1998) stated that such practice is, "consonant with the founding principles of the American university" (p.557). The process of experiential learning, caring for the community, and practicing democratic citizenship are all elements of service-learning methods.

Dewey's (1933) theories of authentic learning were the foundation for Piaget's reflective thinking and cognitive development. Paulo Freire's (1970) authentic learning
and teaching methods also mirrored authentic learning theories, and called for active learning as opposed to teachers merely depositing information. More recently, Koh’s (1984) perceiving and processing premises and Lewin’s (1984) collective and reflective inquiry hypothesis were built on the premise of experiential learning and reflection with action for authentic experiences and education.

These theories have been further expanded by the personal development and empowerment beliefs underscored by contemporary researchers Billig & Furco (Billig & Furco, 2000a; 2000b; Billig, 2000; Billig, 2004). Billig (2002, 2004) contends that service-learning will assist with positive effect on the personal development of students’ interpersonal development as well as acquisition of academic skills and knowledge. Further discussion of these theories can be found in the literature review. However, with the conceptual reasoning of service-learning as a method established historically, the question remains; does service-learning truly impact academic achievement?

Statement of the Problem

Service-learning proponents have asked for more analysis of the impact of service-learning on student achievement on national tests. In “A Proposal to the Field,” Billig and Furco (2002a, 2002b) listed recommendations for future research based on a recent K-12 meeting. In order of priority, their first research question was, “What is the impact of service-learning on academic achievement?” (p. 272).

Billig and Furco (2002b) argue that while a plethora of college-related research exists, very little data exists for K-12. In 2000, Billig published an article synthesizing all of the existing research on K-12 service-learning. Billig documented a variety of studies that illustrated positive impacts of service-learning on student academic learning, school
environment, and the local community. Billig (2000) also reported on several studies of the impact service-learning had on academic skills and knowledge, but found few solid studies correlating the two.

Melchior (1999) published several reports contending that students participating in service-learning programs in high school showed an increase in measures of school engagement and achievement in mathematics. Morgan (2000) has also documented improvements in standardized scores on students in Indiana Schools. Nevertheless, Billig reiterates that continued research is required in experimental and quasi-experimental formats, which study the impacts of service-learning on the student, the community, and the school as an organization.

This study attempted to determine the impact of service-learning programming in New Jersey public high schools on student academic achievement. The study employed both national and state resources by utilizing the New Jersey Service-Learning Leader School nominations in its design. Self-reporting problems were prevented as these schools were chosen by an intensive application and interview process by a panel from the New Jersey Department of Education. Initially, applications are reviewed at the state level. The state department then forwards their nominations to the Corporation for National and Community Service. A national peer review panel reviews the applications and makes recommendations to the chief executive officer of the Corporation for National and Community Service who announces the winners in April of each year (Corporation for National and Community Service [CNS], 2003).

An additional number of schools were also utilized via the school year 2002-2003 Department of Education Service-Learning Survey. The Department sent out a survey to
every public, charter, and independent school in New Jersey. Over 550 surveys were returned including 196 public and charter high school respondents. These particular surveys were integrated into the study design to include schools that have been identified by the Department of Education as fully implanting service-learning into the school culture (CSL Schools) and schools that were not practicing service-learning (Traditional Schools). Questions used to determine the status of service-learning included: number of participating teachers, classes or department that practice integrating service into the curriculum, time of day that service took place, inclusion of preparation, recognition, evaluation and reflection in the lesson plans, and if 50% or more of the student body was involved.

Of the total population of 271 public and charter high schools in New Jersey, 196 high schools responded to the Department of Education Survey. Of these, 96 schools were identified from the pool respondents for this study. The initial 21 schools were chosen because of their identity as Service Learning Leader Schools. Another 27 schools were added to the pool as Community Service-Learning Schools because of their demonstrated commitment to including Community Service-Learning activities as part of the curriculum. This information was collected via the Department of Education Service-Learning Survey.

The remaining schools that did not participate in service-learning school programs were placed in a pool to be matched to the 48 service-learning schools by District Factor Grouping and county. These schools are identified herein as Traditional Schools. The total number of schools used in the study was 78 schools. From the State Report Card Database, information regarding HSPA and SAT scores was gathered for each school. In
some cases, technical schools did not provide test data because the students were tested in
the schools of their sending districts and thus these schools were removed from the
sample. One charter school was closed while the study was being performed and thus that
school and the match were removed as well. The researcher gathered variables including
HSPA math and verbal proficiency scores, SAT math and verbal averages. The end result
was a sample of 78 schools.

In addition to test scores, the researcher gathered information on student mobility
rate, and the percentages of students who participated in testing from the School Report
Card database. This information was gathered for the Service-Learning Leader,
Community Service-Learning, and Traditional Schools and entered into an excel
spreadsheet.

This study utilized 10 null hypotheses to analyze the research data. Three
statistical methods were used to analyze the data: t-tests, Analyses of Variance, and
Analysis of Covariance. A t-test is commonly performed to determine homogeneity of
variance as the test is used to examine if the standard deviations of two populations are
equal (Witte & Witte, 2001, p. 366). In this study, independent sample t-tests were
performed to determine whether the matches were equivalent by school demographics. A
significant mean difference between matched schools for District Factor Grouping,
student mobility rate, or percentage of participation in the SAT would denote an
inequality in the school match and signify a non-homogenous sample.

An Analyses of Co-Variance (ANCOVA) was used to analyze the main effect of
Service-Learning programming and any relationships between Service-Learning
programming and achievement scores while applying statistical control to the variables of
District Factor Grouping (DFG), class size, mobility, and percentages of students tested on the SAT test. These analyses will be discussed further in Chapter Three.

Purpose of the Study

The purpose of this study was to examine how service-learning programming affects academic achievement of New Jersey High School students' tests scores on the High School Proficiency Achievement Test (HSPA), and SAT math and verbal scores.

This study also examined such variables as: student mobility rate, District Factor Grouping, and percentage of students participating in SAT testing. Current research has identified these factors as possible explanatory variables for student achievement. High student mobility has been linked to low achievement (Haveman & Wolfe, 1994; Rumberger & Larson, 1998; Swanson & Schneider, 1999; Teachman, Paasch, & Carver, 1996; Tucker, Marx, & Long, 1998) and current literature has demonstrated that SES is a contributing factor on student achievement and on issues related to the achievement gap (Howley & Bickel, 2002; Strickland & Alvermann, 2004).

Practitioners from the field have reported that much of the existing research concerning service-learning is self-reported from directors seeking grant evaluations. It is not always clear whether such data can be validated (Billig & Furco, 2002a; Billig & Furco, 2002b; Hecht, 2003; Howard, 2003; Melchior & Balis, 2003; Shumer, 2000). Service-learning researchers have recently called for increased rigorous and replicable research in an effort to create a strong platform to advocate for more service-learning programs (Billig, & 2002a; Furco, 2002; Hecht, 2003). While it is firmly believed that service-learning programs impact concepts of self-value, social justice, and community awareness (Billig & Conrad, 1997; Melchior, 1999; Morgan & Streb, 1999), measurable
improvements in student achievement is far more interesting to administrators and school board members who must allocate spending based on the value of each program.

Significance of the Problem

Billig (2002, 2003, 2004) reports that although more than a decade of research has been compiled that focuses on colleges and universities that have institutionalized service-learning, only a paucity of research exists on the K-12 level.

In New Jersey, service-learning has developed significantly over the past decade. Various AmeriCorps projects (Clean Ocean Action, NJ Farmers Against Hunger) have developed a strong presence over the past several years. A long-standing statewide organization, Youth Service New Jersey has continued to hold annual service-learning conferences. The Department of Education itself has supported and funded more than twenty Learn and Serve Grant Programs annually across the state including Abbott districts.

At the college level, New Jersey is home to the Bonner Foundation; a non-profit that is dedicated to creating service opportunities at colleges nationally. The Citizen and Service Education program (CASE) of Rutgers is a nationally recognized service-learning center. Other colleges and universities such as Princeton, Rider, Monmouth, and Montclair have incorporated service-learning into their courses and have hosted annual conferences to further the field.

However, even with such active infrastructure, this is the first statewide study of service-learning programs at the secondary level in New Jersey. This study attempted to discover on a small scale, the potential impact service-learning programming has on academic achievement and thus whether it is as an effective educational pedagogy.
Research Questions

The research question guiding this study was as follows:

How do students’ scores on standardized tests compare between New Jersey high schools that operate state-recognized service-learning programs and those that do not?

The research question was guided by the following subsidiary questions:

1. How do New Jersey Leader and Community Service-Learning School students’ HSPA test scores compare to traditional New Jersey high school students’ HSPA scores?

2. How do New Jersey Leader and Community Service-Learning School students’ SAT test scores compare to traditional New Jersey high school students’ SAT test scores?

3. Are there differences when comparing HSPA and SAT scores of do New Jersey Leader and Community Service-Learning Schools?

For each of the questions the null hypothesis is that there will be no difference between student performance in a leader school and traditional high school was tested. Thus the following null hypotheses was tested in this study. Two null hypotheses were tested for subsidiary Question One; three were tested for research Question Two and five null hypotheses for subsidiary Question Three.

Hypotheses

Subsidiary Question One

1. Students in New Jersey Leader and Community Service-Learning schools will perform no differently than students in a traditional high school on the verbal portion of the HSPA test (H: $\mu_1 - \mu_2 = 0$, where $\mu_1 =$ the population HSPA test mean for the New Jersey Leader and Community Service-Learning schools and $\mu_2 =$ the population HSPA test mean the traditional high schools).
2. Students in New Jersey Leader and Community Service-Learning schools will perform no differently than students in a traditional high school on the math portion of the HSPA test (H: \( \mu_1 - \mu_2 = 0 \), where \( \mu_1 \) = the population HSPA test mean for the New Jersey Leader and Community Service-Learning schools and \( \mu_2 \) = the population HSPA test mean the traditional high schools).

Subsidiary Question Two

3. Students in New Jersey Leader and Community Service-Learning schools will perform no differently than students in a traditional high school on the verbal portion of the SAT test (H: \( \mu_1 - \mu_2 = 0 \), where \( \mu_1 \) = the population SAT test mean for the New Jersey Leader and Community Service-Learning schools and \( \mu_2 \) = the population SAT test mean the traditional high schools).

4. Students in New Jersey Leader and Community Service-Learning schools will perform no differently than students in a traditional high school on the math portion of the SAT test (H: \( \mu_1 - \mu_2 = 0 \), where \( \mu_1 \) = the population SAT test mean for the New Jersey Leader and Community Service-Learning schools and \( \mu_2 \) = the population SAT test mean the traditional high schools).

5. Students in Community Service-Learning schools will perform no differently than students in a traditional high school when comparing total scores on the SAT test. (H: \( \mu_1 - \mu_2 = 0 \), where \( \mu_1 \) = the SAT test mean for the Community Service-Learning schools and \( \mu_2 \) = the population SAT test mean for the traditional high schools).

Subsidiary Question Three

For the third subsidiary question, five separate null hypotheses were tested:
6. Students in a New Jersey Leader school perform no differently than students in a Non-Leader high school on the verbal portion of the HSPA test. (H: \( \mu_1 - \mu_2 = 0 \), where \( \mu_1 \) = the population HSPA test mean for the New Jersey Leader schools and \( \mu_2 \) = the population HSPA test mean for the Non-Leader high schools).

7. Students in a New Jersey Leader school perform no differently than students in a Non-Leader high school on the math portion of the HSPA test. (H: \( \mu_1 - \mu_2 = 0 \), where \( \mu_1 \) = the population HSPA test mean for the New Jersey Leader schools and \( \mu_2 \) = the population HSPA test mean for the Non-Leader high schools).

8. Students in a New Jersey Leader school perform no differently than students in a Non-Leader high school on the verbal portion of SAT test (H: \( \mu_1 - \mu_2 = 0 \), where \( \mu_1 \) = the population SAT test mean for the New Jersey Leader schools and \( \mu_2 \) = the population SAT test mean for the Non-Leader high schools).

9. Students in a New Jersey Leader school perform no differently than students in a Non-Leader high school on the math portion of SAT test (H: \( \mu_1 - \mu_2 = 0 \), where \( \mu_1 \) = the population SAT test mean for the New Jersey Leader schools and \( \mu_2 \) = the population SAT test mean for the Non-Leader high schools).

Students in New Jersey Leader schools will perform no differently than students in a traditional high school when comparing total scores on the SAT test. (H: \( \mu_1 - \mu_2 = 0 \), where \( \mu_1 \) = the population SAT test mean for the New Jersey Leader schools and \( \mu_2 \) = the population SAT mean for the traditional high schools).

Definition of Terms

For the purpose of this study, the following definitions were used.
Service-Learning: Defined by the New Jersey Department of Education as programming within the school that combines community service objectives with the New Jersey Core Curriculum Content Standards. These activities are embedded in the curriculum and take place during the day and include four essential components: preparation, action, reflection and recognition (New Jersey Department of Education, 2000).

National Service-Learning Leader School Program: A national program sponsored by the Corporation for National and Community Service. To date, 200 middle schools and high schools have earned designation as Leader Schools. These schools serve as national models for their integration of service-learning into the curriculum and the life of the school. This initiative seeks to encourage and increase service-learning opportunities for America's students through advocacy, mentoring and recognition (Corporation for National Service, 2003).

New Jersey Service-Learning Leader School (NJ Leader): A school that is part of the national initiative that recognizes high schools and middle schools for their excellence in service-learning. New Jersey nominates twenty schools annually including both middle and high school. Potential schools can be public, Catholic, private, or charter (Corporation for National Service, 2003).

Department of Education Service-Learning Survey: A survey disseminated statewide to Public, Private, Catholic and Charter schools in September 2002. Over 560 elementary, middle, special, and high school districts responded. The survey attempted to collect contact and program information on service-learning programs and included questions that recorded the amount of students engaged in service-learning. The results
were calculated in January 2003 and are available at the Department of State Office.

Community Service-Learning School (CSL): A New Jersey public high school that has been determined to be a school that is fully implementing service-learning programming based on responses on the Department of Education’s Survey on Service-Learning. Questions used to determine the status of service-learning included: number of participating teachers, classes or department that practice integrating service into the curriculum, time of day that service took place, inclusion of preparation, recognition, evaluation and reflection in the lesson plans, and if 50% or more of the student body was involved.

New Jersey High School Proficiency Assessment (HSPA): A state test given to students in the eleventh grade to measure whether they have gained the knowledge and skills in Mathematics and Language Arts Literacy identified in the Core Curriculum Content Standards. The Core Curriculum Content Standards identify what students should know and be able to do at the end of the fourth, eighth, and twelfth grades. The HSPA replaces the Grade 11 High School Proficiency Test (HSPT11), which was administered from 1993 to 2001. The HSPA is a requirement for high school graduation (New Jersey Department of Education, 2003).

Standardized Aptitude Test (SAT): First developed in 1926, the SAT is a nationally recognized examination administered by the College Board. The SAT test measures verbal and mathematical reasoning skills. The SAT math section covers arithmetic, algebra, geometry, probability, and requires students to apply mathematical concepts and use data literacy skills. SAT verbal sections emphasize reading with more than half of the verbal portion devoted to passage reading questions. Vocabulary is also
assessed in sentence completion and analogy questions (College Board, 2004).

District Factor Grouping (DFG): The New Jersey Department of Education introduced the District Factor Grouping system in 1975. This system provides a means of ranking school districts in New Jersey by their socioeconomic status (New Jersey Department of Education, 2004).

Abbott Districts: The Abbott districts are a group of 30 New Jersey school districts that based on socio-economic indicators, are considered to be the poorest in the state. The present Abbott designation of schools came about as a result of the 1985 Abbott v. Burke (100 N.J. 269, 495 A 2d 376) case in which poor urban districts challenged the State's school funding formula. In a series of decisions, the New Jersey Supreme Court granted the Commissioner of Education the authority to ensure adequate funding for poor districts while implementing remedial measures to improve student achievement. These districts are also referred to as "Special Needs Districts" (Liberra, 2002).

State Report Card: In 1995, the New Jersey Legislature passed a law that mandated a New Jersey Report Card for every school in the state. Over the last seven years, the New Jersey Department of Education has expanded the data in the report card to provide as much statistical and demographic information as possible to enable parents and community members to make judgments about the effectiveness of their schools (New Jersey Department of Education, 2002).

Student Mobility Rate: The student mobility rate found on the New Jersey State report card is defined as the percentage of students who entered and left during the school year. The calculation is derived by the sum of students entering and leaving after the October enrollment count divided by the total enrollment. The New Jersey State average
of student mobility at the high school level is 10.9 (New Jersey Department of Education, 2002).

Percentage of students participating in the SAT: The SAT is a voluntary test administered by the College Board, usually for the purpose of college admission. Approximately 76% of each high school’s eligible students participate annually in New Jersey (New Jersey Department of Education, 2002).

Limitations

This study was limited by several factors. One limitation was the size of the study. Only public and charter high schools were compared. These schools are regulated by the State and therefore have a similar makeup of curriculum and student bodies. This study focused on high school statistics only. While Melcher (1999) and Morgan (2000) have found service-learning programming to have a lasting impact in high school students, they did not find similar effects at the middle school level.

For purposes of this study, the research accepted certain delimitations. The data collection was confined to New Jersey Public High School students that have participated in the HSPA and pre-determined SAT scores. Although this study attempted to control for socio-economic factors through District Factor Grouping, variables such as gender, outside coaching and tutoring were not controlled for. Furthermore, the HSPA and SAT scores are not the only predictor of student achievement.

This study was a causal-comparative study. A random assignment of groups was not utilized because the independent variable was already received. The study has weaknesses in its lack of randomization, manipulation, and control.
Summary

Service-learning as an educational method is a developing pedagogy. Nationally, only a limited number of studies have focused on service-learning’s impact on student academics at the high school level. In New Jersey, no statewide studies have been completed or published.

Few school districts in the state of New Jersey employ service-learning methods district-wide. The researcher focused on schools that have achieved excellence in service-learning as determined by the New Jersey Department of Education Leader School Evaluation and the New Jersey Department of Education Service-Learning survey. The scope of this study was focused only on New Jersey public and charter schools.

In this study, SAT and HSPA scores were the determinants of student achievement. The SAT measures verbal and mathematical reasoning skills. The HSPA test focuses on Language Arts Literacy and Mathematics. This HSPA test is tailored to assess whether eleventh grade students have the knowledge and skills required for graduation.

This study has been designed to examine the impacts of service-learning on academic achievement. No other variables will be studied. Academic achievement can also be affected by curriculum, teacher age and education, and professional development.

Organization of the Study

Chapter 1 presents an overview of the study including research questions, definitions, and null hypothesis. The background or the problem significance of the study
was also discussed. Chapter II will contain a literature review describing recent and seminal research on service-learning. Chapter III will present the methodology including a review of the research questions and hypothesis. A discussion of the data collection and population and sample will follow. The validity and reliability of the methodology will be addressed. A summary conclusion will follow. Chapter IV will present the results of the study, and Chapter V will analyze the results and present conclusions.
CHAPTER II
LITERATURE REVIEW

Historical Overview

Service-learning has deep roots in American tradition. Alexis de Tocqueville (1835) first noted the community-oriented nature in his work Democracy in America. Tocqueville discussed the “consistent formation of associations” by Americans, “to further their own needs or combine needs and desires of communities” (de Tocqueville, p.213).

John Dewey (1916) articulated the key principles of progressive education and the need for citizenship education. He advocated for education that focused on experience and action. Dewey believed in the process of active learning. He believed that learning was an inherently social process, which required an increase interaction among students as well as between students and teachers.

Dewey (1933) proposed that pedagogy be recognized as a university discipline. He also argued that academic study was arbitrarily designed, and thus estranged from the contemporary concerns of society. He cited the failure to integrate disciplines and the lack of guided focus of student inquiry on problems as further hindrances to true learning (Ehlich, 1998). Regarding the education process, Dewey (1978) stated:
Learning in school should be continuous with that of out of school process and should engage students in reaching outside the walls of the school and into the surrounding community. It should focus on problems to be solved and that should be collaborative both among students and between students and faculty. (p. 416)

Dewey’s (1933) work was continued through William Heard Kilpatrick, the creator of the Project-Method teaching technique. Kilpatrick argued that real learning should take place outside of school walls and that this process should strive to meet real community needs (Kilpatrick 1918, 1925). It is, “solving problems within a meaningful social context”, Kilpatrick (1932) contended, “that knowledge is best constructed” (p. 320).

During the Great Depression, the Federal Government put the idea of solving problems within a meaningful social context to the test in order to address the economic distress. In 1933, President Roosevelt created the Civilian Conservation Corps, an initiative to employ thousands of workers throughout America by the Departments of Labor, Agriculture, War, and Interior. Roosevelt’s progressive vision sought an immediate solution for the nation’s massive unemployment linked with the concept of public service. As he stated the project would “pay dividends to the present and future generations” by protecting and improving public lands (Shapiro, 1964, ¶ 7). The philosophy of community service thus received the imprimatur of the Federal government.

Thirty years later, in 1961, President John F. Kennedy established the Peace Corps. In less than six years there were 15,000 volunteers. Today Peace Corps volunteers serve in 135 countries. Services include: working with teachers and parents to
improve the quality and access to education for children, work with communities to protect the local environment and to create economic opportunities, and also on basic projects to keep families healthy and to help them grow more food. "The larger purpose, however, is to work with people in developing countries and to help them take charge of their own futures" (Peace Corps, 2003, http://www.peacecorps.gov).

These important Federal initiatives involving service have influenced the development of community service and service-learning programs in post secondary and secondary education.

The development of the concept of "service-learning" originated in the works of Robert Sigmon and William Ramsey of the Southern Regional Education Board in 1967 (Giles & Eyler, 2002). The underlying goal was to focus on "reciprocity." Simon and Ramsey’s principals stated that those being served would control the service provided, while those who served controlled their learning during the process. Perhaps most importantly, those who were served were to become able to serve themselves. (Stanton, Giles, & Cruz, 1999) The tradition of learning while serving thus became an educational focus.

In 1985, the presidents of Brown, Georgetown, and Stanford Universities along with the Education Commission of the States joined together to form Campus Compact. The Compact served as a coalition for college and universities devoted to developing the values and skills of citizenship through participation in public and community service. During its first 17 years, Campus Compact has implemented a range of initiatives designed to make public and community service a fundamental component of American higher education. As service-learning becomes more customary in education, the
Compact began assisting with petitioning local and national governments to recognize its importance. Today there are more than 900 colleges and universities active in the coalition.

Service-learning advocates grew in strength and number until they were able to urge that service-learning become part of the national educational vernacular. In 1989, President George H.W. Bush created the White House Office of Service and the Points of Light organization. The creation of the office set the stage for 1990 legislation of the National and Community Service Creation Act to be ratified. During the long negotiations at the federal level on establishing a national commission, in 1992, Maryland independently became the first state to mandate service-learning as a high school graduation requirement.

American Opportunities to Learn and Serve, the successor to the Community Service Corporation, the Office of National Service and the White House Office of Community Service, AmeriCorps, and also facilitates Learn & Serve America, which provides funding for K-12 schools implementing and maintaining service-learning programs. Learn and Serve America supports service-learning programs in community organizations and schools from kindergarten through college to meet community needs, while improving their academic skills and learning the habits of good citizenship. Learn and Serve grants are used to create new programs or expand and replicate existing programs, as well as to provide training and development to staff, faculty, and volunteers. Access to these grants fostered the growth of service learning programs throughout America. Presently, one in
three schools can boast of hosting some form of a program, even if only in one classroom for a short time (Billig, 2002).

On July 3, 2003, President Bush signed legislation establishing new accounting guidelines for the National Service Trust. The new legislation enabled the Corporation for National Service to authorize 50,000 more AmeriCorps positions in all service categories. In announcing these grants, the Corporation used the new accounting guidelines contained in the Strengthen AmeriCorps Program Act signed by the President (CNS, 2003).

On February 2, 2004 President George W. Bush proposed a 9% increase in funding for the Corporation for National and Community Service in his 2005 budget request (Scott, 2004). The 2005 budget, covers the fiscal year starting October 1, 2004, and consists of $1.018 billion for the Corporation and its programs. This increase represents a change of $82 million over the fiscal 2004 enacted levels. The 2005 budget supports 600,000 senior volunteers, 75,000 AmeriCorps members, and more than 1.5 million service-learning activities for youth (Scott, 2004, http://www.nationalservice.org).

The increase of federal support is reflected in the quantity of service-learning programs. In 2000, Billig stated that service-learning programs existed in K-12 education in every state and that California had joined Maryland with setting service-learning goals for all students. By 2002, Furco was able to add the state of Pennsylvania, Florida, and Minnesota as establishing service-learning initiatives.

During this growing phase, the definition of service-learning has evolved. First defined by Signor and Ramsey, a variety of official definitions have followed. A commonly cited definition describes service-learning as “an experientially based teaching
methodology, which combines classroom pedagogy with community action as an integrated aspect of the course” (Zlotkowski, 1995, p. 22). Most important to defining and recognizing service-learning programs is that the objectives include the key concepts of academic curriculum and reflection.

**Philosophy of Service-Learning**

Service-learning’s theoretical foundation is grounded in experiential learning. Many practices are anchored by experiential learning including community service, internships and outdoor action activities (Flyer & Giles, 1999; Furco, 2002; Howard, 2003; & Saltmarsh, 1996). Commonly, the educational method of service-learning is confused with the practice of community service. While community service as a practice is closely related and has a long tradition in America as well, the tradition of community service is often linked to probation and discipline related measures (Furco, 2002; Saltmarsh, 1996; Slumer & Cook, 1999; Skinner & Chapman, 1999; Stanton et al., 1999). Furthermore, community service also pertains to service “for” rather than “with” the poor (Flyer & Giles, 1999; Furco, 2002; & Skinner & Chapman, 1999) Community service generally is enacted onto a people or place sometimes without preparation and usually without reflection or academic connections. Furthermore, community service is frequently based on social causes and focused on producing civic and ethical outcomes (Furco, 2002).

Experiential outdoor activities may seem similar to service-learning because both are based on the concept of experiential learning. While both may take place outside of the classroom, and rely heavily on reflection for learning, experiential outdoor activities focus on group building and self-confidence.
Internships are sometimes grouped with service-learning programs as well. While internships are reciprocal, they generally benefit the provider. Most internships involve placing a student in a for-profit work environment. The job placements have a primary focus on learning and career development. The creation of placements and programs are based on industry needs.

Service-learning differs from community service, experiential outdoor activities, and internships primarily because it is based on academic disciplines, and provides reciprocal benefits for the student and the site (clients). The focus of the program is intended to produce academic and civic outcomes (Farco, 2002). Though the differences between all three experiential activities are apparent, Farco (2002) contends that "One of the greatest challenges in the study of service-learning is the absence of a common, universally, accepted definition for the term." Possibly adding to the problem is that in the last ten years, over 200 definitions of service-learning have been published (Jacoby & Associates, 1996). Educators and students can be confused regarding the differences. This hinders service learning's promotion as pedagogy and a generation of studies examining its effectiveness (Farco, 2002).

One of the first steps in creating a larger body of scientific evidence on service-learning is to separate community service statistics from that of service-learning. Many public schools are not aware of the differences between community service and service-learning (Farco, 2002; Pritchard, 2002). Ivan Pritchard (2002), found in his evaluation of surveys of public, private, and religious schools that although in many cases the percentage of service-learning programs were decreased, schools were unsure of the difference between community service and service-learning. On some surveys, without
definition of what constituted service-learning programs, schools simply did not report at all (Pritchard, 2002, p. 8). In one report, 75% of the schools reported community service and nine service-learning. However, a closer look at programming revealed that 62% of the schools offered service-learning components that were integrated into the curriculum. While academics, researchers and theorists may favor the phrase service-learning, practitioners in the educational field may recognize and prefer community service as a term for defining their programs and schools.

Public Sector Consultants Inc. (2002) conducted a report for the Michigan Community Service Commission that showed that community service was more prevalent than service-learning with 88% of Michigan public schools participating in community service and 39% of schools in service-learning. Furthermore, it was reported that schools did not differentiate between community service and service-learning, and most of the service-learning taking place was an activity of an individual teacher within the school. The study identified service-learning as structured within an academic course or curriculum with clearly defined objectives and community service as not curriculum based. The results (n = 622) were stratified by level of instruction, with 30% being from high schools.

The Corporation for National Service provides their own definition of service-learning. The definition focuses on four essential ingredients: preparation, action, reflection, and recognition. The preparation through the curriculum combined with reflection guided by educators is the essence that sets service-learning apart from community service. In the National and Community Service Trust Act of 1993, service-learning is defined as:
This definition helps clarify what is central to service learning. Preparation, Action and Reflection are essential for authentic learning to occur. While it is important to reiterate the focus on academic curriculum, it is equally important to include strong components of reflection. These are the characteristics of service-learning that assist students with the authentic learning that educators seek. A long history of research points to the fact that cognitive learning has been facilitated by active and reflective thought.

Cognitive Thinking and Service-Learning

The Corporation for National Service’s definition of service-learning is anchored by Dewey’s belief that genuine learning occurs when lessons and activities are focused on problems in the community. In the 19th century, educator John Dewey embarked upon a lifetime journey to define a new philosophy of education. His efforts became the foundation of what modern service-learning practitioners advocate (Deans, 1999; Ehrlich, 1998; Giles & Eyler, 2002; & Saltmarsh, 1996). Like Dewey, service-learning proponents stress that action and reflection are necessary to foster citizenship and societal
values that will allow democracy to flourish (Barber, 1992). As a "Democratic pragmatist," Dewey's point of view defines the aim of education not in making a student very knowledgeable in a subject but creating a context for the learning experience. In *My Pedagogical Creed* he states, "Unless education has some frame of reference it is bound to be aimless and lacking a unified objective" (Dewey, 1916).

For Dewey pedagogy and epistemology are intrinsically related. His theory of knowledge resulted from his theory of citizenship and democracy (Giles & Eyler, 2002). A central question Dewey (1938) poses is "How are experiences educative?". It is important to remember that not all experiences are educative, or equally educative. Dewey (1938) defines two facets of an experience to gauge whether it is educative: Agreableness and the potential effects for future experiences. If the experience being examined only contains one aspect and not the other, it may have been pleasant (or not) but not educative.

Dewey (1916) challenges educators to play a central role in "saturating the student with the spirit of service" (p.29). Dewey (1933) also stated:

the growth of the child in the direction of social capacity and service, his larger and more vital union with life, becomes the unifying aim of [of education]; and discipline, culture, and information fall into places as phases of this growth.

(p.92)

Dewey's (1938) theory of learning is comprised of two principles: continuity and interaction. The principle of continuity posits that all experience occurs along a gamut called the experiential continuum. Throughout life, experiences build upon the previous ones. These experiences, however, need direction and development provided by the
teacher. The second principle is one of interaction in which aspects of the experience interact to form the situation at hand (p.42). Learning results from the transaction between the individual (learner) and the environment. These two principals interact and unite to form the "longitudinal and lateral aspects of experience" (Dewey, 1928, p.44).

The principals of continuity and interaction incorporate both reflective thinking and action. Dewey's (1933) definition of reflective thinking is, "Active persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusions to which it tends" (p. 9). The term "reflective thinking" is used to indicate a type of thinking and the term "reflective activity" to identify the complete set of activities related to the reflection. What is central is that thinking and action are inextricably linked (Eyler & Giles, 1999).

Dewey (1933) defined reflective thought as having five phases or aspects. Although he lists them chronologically, this is not a linear or cumulative process. Dewey specifically wrote that the "sequence of the five phases is not fixed," and could be collapsed or expanded, but they were "the indispensable traits of reflective thinking" (p. 115-116).

1. Suggestions – The inhibition of tendency to act, to pursue what suggestion arises from the situation

2. Intellectualization – definition of a problem and the raising of questions about the problem and possible solutions

3. The hypothesis – this is the development of the guiding idea based on observation and previous knowledge
4. Reasoning - this is the development of the hypothesis by applying knowledge and by developing the linkages in the sequence of ideas.

5. Testing the hypothesis in action - this is the verification through further observation or experimentation in which the problem is solved or a new problem is presented. (Dewey, 1933, p. 10-175).

While Dewey defines ideal dialectic in terms of continuity and a move to civic participation, Freire uses the term "dialogue" and hopes for critical consciousness (Deans, 1999, p. 20). Self-labeled progressive and radical, Brazilian educator and advocate Paulo Freire became renowned for accounts of his methodology and experiences in teaching Latin American peasants how to read. Freire (1970) utilized his own terminology for processes he created and encountered during his tenure in teaching. The process of "banking" refers to inserting, reiterating, or merely depositing the information (lesson) in the students' head (rote memorization). Banking, according to Freire, only produces injustice by indoctrinating the student to passivity. Freire (1970) stated, "In the Banking concept of education, knowledge is a gift bestowed by those who consider themselves knowledgeable upon those whom they consider to know nothing" (p. 53).

Conversely, a critical education assists the students in methods to understand the world (p. 72). According to Freire (1970), two traditional dichotomies within education; separation of teacher and student and the resulting separation of action and reflection are avoided with genuine learning. This philosophy translates easily into service-learning. In unifying the traditional roles and incorporating the Freire-defined acts of reflection and action, all involved in the education process gain power. In doing this, Freire states that
the process will create human beings, rather than support roles that reinforce an
oppressive system (p. 115).

Freire's writings are heavily influenced by his study of third-world literacy,
having spent much of his lifetime focusing on the marginalized and dispossessed (Dean, 1999). For Freire, action and reflection combined together is what essentially makes
people human. He states, "Action is human only when it is not combined together, but
also a preoccupation, that is, when it is not dichotomized from reflection" (Freire, 1970,
p. 35). In addition he states, "Authentic reflection considers people in their relations with
the world" (Freire, 1970, p. 73).

Like other educational progressives, Freire stressed that the heart of all learning
rests in the way we process experience. Of particular importance is our critical reflection
of experience. Learning is more a cycle beginning with the experience, continuing to
reflection, which can later lead to action, which itself becomes a concrete experience for
reflection (p. 84).

Another term created and utilized by Freire (1970) is "praxis," which he defined
simply as "action-reflection." He asserts, "Within the word we find two dimensions:
reflection and action, in such radical interconnection that if one is sacrificed—even in part
the other immediately suffers" (Freire, 1970, p. 75). Freire believes that to reflect
without action leads to "verbalism" or empty words. Moreover, action sans reflection
leads to activism without critical analysis. This, according to Freire, is essentially
unauthentic and reinforces the original dichotomy as he states, "Authentic reflection
clarifies future action, which in its given time will have to be open to renewed reflection"
(Freire 1985, p.156). Freire maintains that true education, "consists in acts of cognition,
not transfers of information." Similar to Dewey's (1938) view on stagnant disconnected classroom learning, Freire envisions active student interaction with the world (p. 164).

Jean Piaget, a Swiss philosopher and psychologist, proposed a model of cognitive development that complements the importance of direct experience. Contrary to the behaviorists who view humans as passive respondents to the environment, Piaget (1952) argues that people are active processors of information. It is through this interaction with their physical and social environment that children learn and develop cognitively.

Piaget's focus is on cognitive-development processes, on the nature of intelligence, and how it develops. In his study of the development of cognitive processes, Piaget believed childhood to be the key to understanding the nature of human knowledge itself. Piaget's theory describes how intelligence is shaped by experience. Intelligence is not an innate internal characteristic of the individual but arises as a product of the interaction between the person and his or her environment. For Piaget, action is key (Kolb, 1984).

Piaget's theory of cognition contains three basic principles: assimilation, accommodation, and reflective abstraction. The principle of assimilation, is defined as where the mind primarily assimilates, perceiving and categorizing experience in terms that are already known. The second principle of accommodation is the mechanism for learning. If during the initial processing phase, the mind finds an experience that does not fit into an already existing category, an evaluation begins. This evaluation can potentially lead to a change in an existing structure or the formation of a new one; thus, an accommodation (Piaget, 1952). Piaget (1952) describes intelligent adaptation as the result
of a balanced tension between these two processes. If the balance is upset
and accommodation processes dominate assimilation, we have imitation, which Piaget
describes as, "the molding of oneself to environmental contours or constraints."
Furthermore, when assimilation predominates over accommodation we have "play"
which is defined as, "the imposition of one's concept and images without regard to
environmental realities." The process Piaget describes is of cognitive growth from the
concrete to abstract. It is a continuous operation between assimilation and
accommodation. Along with the process of action and reflection, cognitive thinking
continues in successive stages, each higher than the previous (Piaget, 1952).

In Piaget's theorem, the third principle is deemed "reflective abstraction". This
principle pertains to reflection on mental operations, and provides a model for how it
operates. In his view, the organization of immediate experience, the "sensorimotor
intelligence" that exists in simple action schemes, and the basic ability to consider one
thing as the symbolic substitute for another, are cognitive functions of the child before it
has any conception of other people, let alone their common social practices (Piaget,
1952).

Piaget describes the cognitive development of children in four stages:
sensorimotor, preoperational, concrete operations, and formal operational stage. It is
important to note that the formal operational stage, the last of Piaget's developmental
stages, supports thinking that is abstract, formal, and logical. In this stage, thinking
includes events that cannot be observed. According to Piaget, while a child may be able
to think hypothetically and thus use logic to solve problems, not all individuals reach this
stage. Cognitive studies have reported that only 40-60% of college students fully achieve
this stage of development and in some undeveloped countries almost none have (Kodat, 2002). Contemporary cognitive scientists continue to research reflective thinking in an effort to address this disparity (Eyler, 2002).

Kolb (1984) is a cognitive scientist whose work has advanced the concept of reflection and cognitive learning by dividing it into two separate learning activities: the perceiving and processing stage and the abstract conceptualization stage. Kolb describes this process of moving back and forth from experienced thought as learners assemble and organize knowledge. His model begins with concrete experience, followed by reflective observation, then abstract conceptualization, ending with active experimentation in which questions are pursued or the theory is tested through more experience.

Kolb (1984) attempted to illustrate the reflective process through a certain pattern. While in the critical reflection stage, he contends, we may ask questions about the experience in terms of previous experiences. In the abstract conceptualization stage, we try to find answers, making generalizations, and draw conclusions and hypotheses about the experience. The Action phase is the phase of active experimentation, where we try the hypotheses out, revise and try again. Of this phase, one would typically rely on systematic planning of theories and ideas to solve problems as Kolb (1984) stated, "In this stage, learning involves using logic and ideas, rather than feelings to understand problems or situations" (p. 74).

The Kolb model of the continuous learning cycle is used frequently in modern pedagogy since it provides a foundation for creating longitudinal lesson plans (Eyler, 2002). While Piaget argues the notion that the capacity for abstract thought develops over
time, Kolb stresses the continuous movement back and forth between the concrete and the abstract which creates a learning cycle.

Kurt Lewin was a seminal theorist who deepened the understanding of groups, experiential learning, and action research (Kolb, 1984). While Lewin is best known for his leadership theory, his research and discovery of T-Groups are the basis for team building in organizations. T-Groups are defined as a tool that facilitates learning by bringing together a small group of people for the express purpose of studying their own behavior when they interact within a small group. This grouping provides participants with an opportunity to learn about themselves and how to function more effectively in group and situations. Through the T-Group innovations, Lewin documented that learning is best facilitated in an environment where there is "dialectic tension and conflict between immediate concrete experience and analytic detachment" (Kolb, 1984, p. 10).

Lewin's second discovery was borrowed from the field of electrical engineering. "Action Research," a term that he is well known for, was based on the process of "feedback." Action research serves as a social learning and problem-solving process that generates information in order to access and evaluate goals. This activity takes the form of collective and reflective enquiry. The participants must work collaboratively in social situations in order to improve their practices, as well as their understanding of those practices and the situations, in which the practices are carried out (Lewin, 1951). This approach to research is oriented to problem solving and is similar to Dewey's concept of learning from experience.

Leading experiential theorists and cognitive scientists have touched on the important facets of genuine and authentic learning. Common to all is the concept of
active learning. Active learning involves reflection as well reciprocating amongst learners and teachers. The theories expounded by these leaders include an emphasis on creating a better individual and a better society via the education process.

Furco (2002) underscores this conviction in recent studies when he reports that students are empowered by respect and adult responsibilities. When students feel empowered, they enjoy the experience more (Billig & Furco, 2002a, p. 43). Billig (2002) also contends that service learning impacts personal development. Eyster, Giles, & Braxton, 1997) back up these findings in their study of Vanderbilt students.

Service learning meets the parameters of what established research considers being requisites for effective cognitive learning. The action, reflection, and meaningful experience are anchored to the academic curriculum. Experience coupled with the ensuing reflection is the reason why service-learning is such a powerful pedagogic movement.

Experimental Education and Service-Learning

Direct Experience

Direct experience during the “action phase” of a service-learning project is a potent catalyst. Numerous studies in education and social psychology emphasize the importance of the direct experience and how this aspect of learning provides memorable context to the event (Billig, 2002; Blyth et al., 1997; Fazio, 1986; Giles, 1995; Morgan, 1995).

Fazio and William (1986) demonstrated that direct exposure to a topic or attitude would cause related feelings and behaviors to be stronger. Fazio and William reported that the subsequent feelings and behaviors would be more readily accessible. The reported results of the same situations reenacted devoid of direct experience with attitude
objects showed only a weak connection. These findings suggest that because attitudes from direct experiences are more accessible, they are more likely to persist over time.

The contentions advocating direct experience are underscored by the studies that illustrate that direct service benefits personal development (Blyth, Saito, & Berkas, 1997). Giles & Eyler (2002) found that direct service to others could potentially be an effective way to combat stereotypes. Morgan (1995) described similar findings in a high school-based study.

Often community service programs are facilitated without the participants witnessing what they are working towards. Indirect service such as working at a food pantry on a day without clients, raising money for an international cause, or donating food cans or clothing in a collection box. In all of these cases, the students or servers do not get to experience the direct effects of their efforts. Direct experience however, provides the student or server with a personalized experience and a context for learning.

Direct experience and direct service is only a piece of the service-learning matrix. The experience itself will provide a potentially long lasting memory. However, attitude and behavior need to be critically analyzed to gain any true understanding of the experience. That is why reflection is such a critical element of the service-learning process.

Reflection

Drawing from the work of experiential and cognitive theorists discussed previously, a case can be made regarding why reflection is often cited as one of the most important factors of a service-learning (Eyler, 2002; Fenzel & Leary, 1997; Gray et al., 1998; Mabry, 1998; Melchior, 1999; Saltmarsh, '996). Conrad and Hedin (1991) found
that opportunity for frequent reflection was the best predictor of outcomes. Studies which have examined the impact of quality differences in service-learning have found that programs with more opportunity for reflection, substantive links between course work and service, and ethnic and cultural diversity have a stronger impact (Eyler & Giles, 1999; Gray et al., 1998; Mabry, 1998).

Dewey (1916) understood that, "When we reflect upon an experience instead of just having it, we inevitably distinguish between our own attitude and the objects toward which we sustain the attitude" (p.173). Without reflection following an activity, the connection between thought and action dissipates and the ability to formulate further action (reflective action) is lost.

In a study based at a parochial college, Fenzel & Leary (1997) concluded that reflection needed to be a central feature of a program. The study focused on philosophy classes and included 134 students in six different class levels. The central feature of the classes was reflection to stimulate students’ thinking about issues of justice and faith. Findings supported the conclusion that reflection was essential.

Mabry (1998) evaluated how students' attitudes and values were affected by the amount of service and frequency of reflection. The findings published concluded that service-learning was effective only when students had at least 15-20 hours of service, engaged in weekly in-class reflection, and wrote reflective assignments.

Gray et al. (1998) published the RAND study of 500 higher education institutions and community organizations during 1995-1997 fiscal years. Research methods included surveys, site visits, and data analysis. Over 1,306 students from 28 institutions completed surveys. Results included the finding that in order for positive benefits to last, service
experiences must be for an extended amount of time and that it is important to link service-learning with course content. Reflection continued to be a critical component of service-learning in the form of class discussions, journal writing, essay, or research papers.

The results concluded from surveys of 1,500 students from 20 colleges and universities by Eyler et al. (1997) also supported reflection. The report concluded with the premise that reflection can increase impact of the service on the students as well as improve the relationships between the students and educators. The study also reported service-learning to be a predictive of a career of valuing people, volunteering, and political participation.

Ikeda (2000) researched the importance of structured reflection in service learning courses. The qualitative case study included interviews with 15 faculty and 72 students from three colleges. Results included the statement that structured reflection in service-learning courses was the only method teachers used to connect the service experience to the academic course. The types of reflection included: class discussions, written journals, and papers. Ikeda concluded that reflection is essential to provide structured, intentional contemplation and enhances learning by connecting service to academic courses.

Various studies have supported the contention that reflection plays a central role in the process of learning (Eyler & Giles, 1999; Eyler, Giles, & Braxton, 1997; Waterman, 1997). Reflection provides the transformative link between the action of serving and the ideas and understanding of learning (Eyler et al., 1996, p. 14). “Service combined with learning, adds value to each and transforms both” (Kendall, 1991, p. 95).
Service-learning has the potential to be an effective educational technique, and a powerful movement. However, a major hindrance to its evolution and its broader acceptance by educators is the lack of organized research. A 1991 synthesis of literature by Conrad and Hedlin declared that without quality research the case for service-learning is relatively "unproven." Presently, this problem still exists (Billig, 2000; Furco, 2002; Prichard, 2002). The thrust of the research focusing on service-learning is at the college level and primarily draws from qualitative interviews and case studies.

Service-Learning Research

College-Based Research

Marcus, Howard, and King (1993) conducted one of the few studies that attempted to isolate the effects of service-learning on academic achievement (Bersot, Younkin, & William, 1998). The research was an experimental study utilizing eight sections of an American Politics course. Randomized control group design was employed as two sections did not use service-learning and were compared to six that did. The research plan controlled for achievement, because students were registered for classes before the sections were chosen. Marcus et al. found higher scores on mid-term and final examinations and concluded that the grades in a randomly selected group of students were better (blind grading). The service-learning students reported high satisfaction in course evaluation, and awareness of society and societal problems. Similar results were found with a repeated study. The potential flaw was that while lectures and exams were consistent across classes, different individuals facilitated discussion sessions.

In 1997, Astin and Sax surveyed 3,450 students from 42 college institutions. Their findings revealed that the impact of service-learning on a student increased as time
devoted to service increased. They found that service-learning positively affected the student's grade point average and graduation rate. However, as noted in the report, their study lacked a control group. This study was followed in 1998 with results from a quasi-experimental study of 12 colleges and universities. Utilizing data from a freshman survey, SAT and ACT scores, and enrollment data, Astin and Sax tested for 35 student outcomes. These results included the academic outcomes; grade point average, retention, degree completion, and increase in knowledge were favorably influenced by service.

A larger longitudinal study was conducted in 1999 when Astin, Sax, and Avalos (1999) published findings from an original sample of 27,000 students. The research method was to survey students beginning in 1985 with the second survey following four years later. A final survey given to 12,376 students (of the original 27,000) followed in 10 years. The report included a strong correlation between volunteering in high school and as an adult. Behavior and demographic variables such as highest degree earned, degree aspirations, and attending graduate school were associated with the volunteering variable.

Berson, et al. (1998) studied the effects of service-learning on student success in college. Their study consisted of 286 students enrolled in six matched community college courses. One section of each pair was taught using the traditional syllabi, while the other section of each pair included the requirement of service-learning placement. The results indicated that students who participated in the section with a service-learning requirement achieved higher final course grades and reported greater satisfaction with the course. The reported mean final course grades were .26 higher for the students in the treatment group (p<.05).
Eyler and Giles (1999) facilitated a quasi-experimental design with a control group. They compared 1,100 college students participating in service-learning projects to 400 students who were not. The reported results included positive effects on personal developments and interpersonal skills. Other results included significant changes in the complexity of the students' problem analysis skills. Eyler and Giles concluded through the data collected on the reflection aspect of the students' service that students generally remembered and enjoyed their experiences interacting with other people most (p.172). The study did not include an instrument to report independent effects for race, but did include a control group. The self-selection aspect of the study is important to note, because students were able to choose to be in the service-learning class as opposed to the traditional course, this may have drawn students who already exhibited the positive characteristics that were reported as being an effect of service-learning.

Eyler and Giles then continued research with three studies assessing impact of service-learning. The first study was a survey of 1,500 (400 control) students from 20 colleges and universities. The reported results showed that service learning had a positive impact on outcomes of personal development, social responsibility, and interpersonal skills. Students reported that they were more intellectually stimulated, and worked harder in these courses (Eyler et al. 1997). The majority (80%) of the service-learning students reported that they learned more and were motivated to work harder in the class. Eyler and Giles found that well integrated service-learning programs produce greater learning with high quality experiences. With the application of problem analysis it was found that service-learning students were able to create practical strategies and apply subject matter when problem solving.
The second study was a pre and post in depth study with 66 students in service learning courses from 6 colleges. The focus of the study was how service impacted the placement quality, link between academic matter, and reflection. The study found that outcomes were impacted significantly by the quality of the service. Eyler and Giles (1999) followed with a third study attempting to address alternative means of including reflection in a service learning course. This study was sponsored by the Corporation of National Service and included students from 65 students from 6 colleges. Eyler and Giles confirmed that students were able to apply subject matter and experience more easily through structured reflection.

Reeb, Sammon, and Issackson (1999) examined the relationship between service-learning and academic achievement in an abnormal development course. A positive relationship was reported (course examinations). The student self selection aspect, however, needs to be noted.

Siastrī (1998) focused on psychology courses. The research was designed to measure content knowledge gains by 64 students in two class sections. In the study, the instructor and course content remained the same. One section of students enrolled in the service-learning course, and completed a project requiring a minimum of 20 hours. The final grades of the service learning course students were found to be 10 points higher, but not statistically significant. Written assignments were reportedly higher as were the total scores of materials (quizzes, exams, assignments).

Also focusing on the discipline of psychology, Strage (2000) looked at child development courses and utilized the same 20-hour requirement. The study compared student performance of 477 students over five semesters on three examinations (multiple
choice and essay). The students in the service-learning course received a greater number of points on the first exam, but it was not significant. The final two exams and essays reflected higher scores however.

Moely et al. (2002) recently surveyed over 300 college students participating in service-learning and included a control group for both pre and post testing. The results showed that those who were engaged in service-exhibited increases in their plans for future civic action, as well as in their interpersonal, problem solving, and leadership skills.

Reported research continues to build a strong case in favor of the positive effects of college-based service-learning. However, at the high school and middle school level, the curriculum relies upon the leadership of educators and administrators who commonly do not have time to create extra projects much less conduct studies on them. As a result, studies by teachers who actually design and implement service-learning projects have been notably absent (Root, 2003). Consequently, although a great many schools do include service-learning programs within their curriculum very little is known about them (Billig, 2004).

**High School-based Research**

While there have been numerous studies on the impact of service-learning on college students, less have been reported concerning secondary school students. Shumer & Cook (1999) reported that the number of high school students involved in service-learning had increased 3.663% since 1984. This was particularly interesting given that the number of high schools had remained relatively unchanged over the same time period in a nearly static population of 12,377,455 in 1984 and 12,615,913 in 1997. In the study,
Shumer and Cook differentiated between community service and service-learning by the phrase “integration into the curriculum.” They found a 686% increase in community service activities also.

The enactment of federal legislation to support national and community service has proven to be a catalyst for the growth of service-learning programs. The Corporation for National Service (CNS) is the largest single funding source of initiatives in elementary, secondary, and post-secondary institutions promoting service-learning (Corporation for National Service, 2003). Shumer and Cook’s survey results Wite & Wits, 2000, p. the receipt of Learn and Serve America grants from the Corporation for National Service as well as mini-grants to teachers as the most common reason for creation of new service-learning programs.

In 1999, the National Household Education Survey (NHES) released data stating that student participation in service activities had grown to include over half of 6th through 12th grade students. NHES also found that half of these students participated in service-learning (NHES, 1999).

In 2003, the Institute for Social Research of the University of Michigan’s Research Center conducted its annual survey of volunteering. The Monitoring the Future (MTF) data collection provided data from over 50,000 8th, 10th, and 12th graders nationally. Where volunteering rates had remained steady at 67% since 1976, MTF found an increase to 76% in 2001 (Lopez, 2003).

Williams (1993) attempted to determine impacts on suburban high school students. He found that ten hours of service or less created similar socially responsible
attitudes compared to a student who did not perform any service at all. Astin and Sax (1997) found similar results in the study of the impacts on college students.

Follman and Muldoon (1997) were able to utilize a large sample (n = 50,000) of Floridian high school students involved in 197 service-learning projects. Their report for Learn & Serve America revealed improvements in attendance, grade point average, and discipline referrals. At-risk students evidenced a higher percentage of positive increases. They determined that the impact of service-learning was further intensified by group experience (Follman & Muldoon 1997; Morgan, 1995). However, the Learn & Serve data was collected mainly from project directors of Learn & Serve Grants. These directors while more likely to report consistently, are also likely to report positive findings over negative ones in order to receive more funds.

National Center for Educational Statistics reported in their National Student Service-Learning and Community Service Survey (NS&CS) in 1999 that 64% of all public schools, and 83% of all high schools offered or mandated community service. They also found that 53% of all public high schools offered service-learning programs (Skinner & Chapman, 1999).

Melchior (1999), surveyed 17 schools randomly (ten high schools, seven middle schools) with a variety of strong service-learning programs in a study sponsored by Brandeis University for Learn & Serve America. Some of the programs were school-wide, and others classroom or subject specific only. The common variable was that each school had a minimum of sixty hours of service-learning per student in the program (ed to the curriculum. Melchior found an increase in communication skills and found
significant and long-lasting effects on measures of civic participation in high school students. However, no such effects were observed amongst middle school students.

Morgan also reported in 2000 that standardized scores in Indiana elementary schools improved with service-learning participation. His findings included a reported, "robust correlation when controlling for demographic characteristics" (Morgan, 2000, p. 1).

Kraft & Wheeler (2002) ran a small qualitative study in 2002 of the relationship between service-learning and resilience in disadvantaged at-risk high school youth. This study took place at an alternative high school, all of the attendees participated in service-learning. In terms of achievement, students clearly made significant progress in reading and writing and in school adjustment and general resiliency. Student attendance and participation rates increased, as did their grades. Students exhibited an increased positive attitude toward themselves and school. Where half of the student body had been adjudicated, the program was found to engage many of the students in more appropriate behavior, better academic achievement and school attendance. The group studied however, was quite small (n = 39), and without a control group.

For the amount of service-learning taking place at high schools around the United States, there is a scarcity of large scale or multi-program research focusing on its existence and impact (Billig, 2003). Conversely, the increase in service-learning programs in the absence of supporting evidence is noteworthy because it has occurred during a time of increased emphasis on assessment and accountability in all education (Bringle & Hatcher, 2000). The majority of studies are at the secondary level and surveys
of the self-reporting nature. Unfortunately, self-report surveying is not only a weak measure of cognitive outcomes; it also confuses satisfaction with learning (Eley, 2002).

Other studies neglect the factor in the issue of selectivity in service-learning programs. Many times in reports of better grades and outcomes in a class, the question remains whether better students may choose to become engaged in service or does service lead to better grades (Astin & Sax, 1997; Melchior, 1999).

While attempts have been made to use grades as measures of learning, the evidence is mixed (Borsom et al., 1998; Kendrick, 1996; Markus et al., 1993; Miller, 1994). In certain studies where positive results are reported the method of calculating grades has differed for treatment and comparison students. For example, results may be attributed to vague standards such as 'extra credit' rather than increased learning (Balsadeh, 1996; Sugar & Livosky, 1998).

In Heads, hearts, and hands: The research on K-12 service-learning, Shelley Billig (2004) compiled the most current research on service-learning for the National Youth Leadership Council Conference. The organizing theme of her summary was that service-learning has an effect on students in three domains: cognitive ("heads"), affective ("hearts"), and behavioral ("hands") along with effects on schools and communities (Billig, p. 10, 2004).

In cognitive service-learning research, there were several new studies discussed. In Michigan, Billig & Klute, (2002) and Klute & Billig, (2002) had organized a study of 1,988 students respondents, with 1,437 of these students as participated in service-learning. The results of the studies showed that grades 7th-12th were more engaged cognitively in English Language than comparison students. In the younger grades of
2nd-5th, a statistically significant difference in all aspects of cognitive development was found. These aspects included student engagement, more likely to pay attention to schoolwork, concentration, and trying as hard as they could on problems. The study was replicated using the MEAP Test for 5th grade statistical tests showed that students involved in service-learning scored significantly higher on the writing test, the social studies strand scores: historical perspective, geographic perspective, and inquiry and decision-making.

Other recent studies have attempted to address cognitive learning and academics not only by way of scores, but also by content standards and attitudes. The California Comparisons Study performed by Farooq, in 2002 defined academic outcomes in terms of mastery of course content, thinking and problem solving skills, and attitudes toward learning. Students engaged in any type of service had significantly higher scores on survey that measured attitude toward school. In this study, the service-learning groups scored higher in all academic measures. New England CO-SEE Sites Analysis showed that sixth grade students had significantly higher achievement scores on state assessments in the areas of language arts, mathematics, science and social studies than past average. No difference was found in the 3rd grade (Klute, 2002).

More recent studies found statistically significant positive impacts in programs in Alternative schools, (Laird & Black, 2002; Kraft & Wheeler, 2002; Billig, et al., 2003) overall participation and perceived learning in public schools,(Kirkham, 2001) and overall learning compared to regular instruction (Billig & Meyer, 2002; Hecht, 2003; Kirkham, 2001; Molchior & Balić, 2002). However, it is important to note the findings of Ammon, et al., (2001) which denote the following factors related to higher academic
impact included: clarity of academic goals, clear connections between goal and activities, reasonable scope and support through reflection activities. Only successfully planned and maintained programs yield positive impacts.

Summary

In today’s educational climate, the synthesis of literature compiled by Conrad and Hofin in 1991 is still relevant (Billig, 2000; Furco, 2002; Pritchard, 2002). Very few of studies involving high school programs utilized control groups or maintained experimental or quasi-experimental designs. Since much of the collected data was from site evaluations, the majority of the research used self-reporting, and case studies. Rarely did the studies include the hypothesis or the theoretical basis that the curriculum operated on (Billig, 2000).

In two recent volumes of service-learning research, Shelly Billig and Andrew Furco include research plans for K-12 service-learning in the future. The field, according to Billig and Furco (2002a; 2002b) is saturated with anecdotal evidence, but lacks experimental and quasi-experimental studies. To continue to support the movement of service-learning and to assist practitioners in the field, research should be undertaken with a more scientific approach and must study quantifiable outcomes.

In both reports, Billig and Furco (2002a; 2002b) discuss the need for practitioners to create, “Replicable studies that underscore the distinct identity of service-learning” (p. 222). A study of the impact of service-learning on student’s academic achievement in New Jersey will fulfill these mandates for scientifically based research while providing quantifiable evidence for educational leaders wanting to implement service-learning.
programs in their school curriculums. Chapter III will elaborate on the research design, methodology and the projected population, sample.
CHAPTER III

METHODOLOGY

This chapter describes the methodology that was used for this study. A review of the problem statement is presented followed by research questions and null hypothesis. The chapter also contains information regarding the population and sample as well as the validity and reliability of the tests used in the study. The data collection and analysis of the data of the study are also discussed.

Research Design.

This study employed a non-experimental cross sectional research design (Johnson, 2000). In a causal-comparative study, the researcher attempts to determine the cause of pre-existing differences in populations (Haller & Kleine, 2001). Because these differences are pre-existing, the study is classified as "Ex post facto" (after the fact), where the effect and alleged cause have occurred and will be studied retrospectively. A causal-comparative design is used to compare groups that differ on some characteristic in order to determine if this difference might be related to important variables of interest (Haller & Kleine). Generally in this analysis, the researcher attempts to determine the cause, or reason, for pre-existing differences in groups of individuals.

For this study, the independent variable is service-learning. High schools included in this study are drawn from two subgroups: they are either recognized by the New Jersey Department of Education as administering top service-learning programs or
not. A further subset of subgroup I will be Service-Learning Leader schools (n=17) and high schools determined by the Department of Education Survey as practicing community service-learning school-wide (n=20). Schools were then further classified based on their District Factor Grouping letter and their school size.

This study also examined the school demographic variables of student mobility rate and district factor grouping as current research has identified these factors as possible explanatory variables for student achievement (Fisher et. al, 1996; Jencks & Phillips, 1998; Skandera & Sousa, 2001). A factor related to successful academic achievement is regular attendance, or stable mobility in the same school (Kerbow, 1996; Temple & Reynolds, 1999). Average achievement scores of schools with a high rate of student mobility are much lower than schools with more stable enrollments (Beck, Kratzer & Isken, 1997; Mao et al., 1998; Neuman, 1987; Vail, 1995). Not only are higher rates of mobility related to lower achievement (e.g., Mao, Whitsett, & Melkka, 1998), mobile students exhibit more behavior problems, academic deficiencies and social inadequacies in comparison to stable students (Alexander, Dauber & Enwistle, 1996; Brown & Orthner, 1990; Ingersoll, Scamman & Eckering, 1989).

Currently, state educational departments mandate standardized testing programs to collect data about student achievement and to hold schools accountable. Schools are judged and rated according to the average scores of their students. These policies were intended to assist with improving education by setting benchmarks for high standards of achievement. The public reporting of test results was also meant to assist with achievement by allowing communities to compare and contrast between districts. The term "high stakes" for achievement testing evolved as the practice of reporting became
commonplace for grade promotion and academic awards as well as financial penalties and student demotion when scores are below standard. This practice has been exacerbated by the current No Child Left Behind legislation which focuses on standardized testing scores as a rating for school-wide achievement. The new legislation also calls for disaggregated scores at all levels of testing. Schools will clearly be held more accountable for low scores, especially those associated a particular race.

With high stakes testing rising in popularity, the researcher has chosen to utilize standardized testing results as a measure of academic achievement. If the results of this study indicate that service-learning programming has a strong impact on academic achievement in districts may choose to adopt policies of implementing more service-learning programming.

The SAT test is the single most-used standardized test of its type in the country (College Board, 2002). The SAT I: Reasoning Test, a multiple-choice exam with verbal and mathematics components, is used in combination with high school grades to predict a student's readiness for college. Mean SAT scores are tracked annually as indicators of the quality of education in school districts, states and nationwide (Bracey, 1996; Powell & Steedman, 1996).

However, when comparing data from different districts, counties, and states it is also important to take the percentage of student population taking the SAT into account. This number represents the percentage of 10th and 11th grade students with high academic aspirations according to a report from The College Board on SAT test takers. In 2003, more students took the SAT than ever before. The 1.4 million SAT takers in 2002-2003 indicates the largest increase in the number of SAT takers in more than 15 years (College
Board, 2004). According to College Board, Thirty-eight percent of those SAT takers are first-generation college-bound students, and Forty-two percent of 2003 college-bound seniors reported grade averages of A+, A, or A-. The average grade point average of SAT participants was 3.29. Thus the percentage of the school population that actively seeks college admission is not only an indicator of achievement but like student mobility rate, could have a potential impact on this study and needs to be considered.

This study will also control for socioeconomic status (SES) via school District Factor Grouping (DFG) assignment. Current literature has demonstrated that SES is a contributing factor on student achievement and on issues related to the achievement gap (Strickland & Alvermann, 2004). Achievement disparities between DFG and race/ethnic groups occur annually on assessments in language arts literacy (NJ Department of Education, 2003). In a study of 13,600 schools and 2,300 districts, it was found that the odds of getting high test scores are improved by high SES and made worse by low SES (Howley & Bickel, 2002).

The dependent variable in this study is student achievement. The 2002-2003 school year mean scores of the verbal and math portions of the SAT and New Jersey MSPA test were collected to be used as a measure student achievement for 2003. Gay & Airasian (2000) describe the source of weakness as a causal-comparative study as being the lack of randomization and manipulation. Solutions ascribed to these weaknesses include: random assignment of participants to groups, pair-wise matching, comparing homogeneous groups and subgroups, factorial analysis of variance, and analysis of covariance (p. 352).
In this study, pair-wise matching and the formation of homogeneous subgroups were employed. This entailed matching and comparing of groups that are homogeneous in order to control for extraneous variables. Because socioeconomic status and school size could likely influence the dependent variable of student achievement, service-learning leader schools and practicing Community Service-Learning (CSL) schools were divided by their DFG assignments. The CSL schools were then matched to a traditional school from the same district, and DFG assignment. If a similar school did not exist in the district, the match was made within the county, still utilizing DFG and size similarity for matching. It is these subgroups within the matching technique that enable the researcher to compare results from different schools that may have previously been limited by their socioeconomic difference (Gay & Airasian, 2000).

Problem Statement

The purpose of this study was to examine how service-learning effects academic achievement of New Jersey High School students’ test scores on the High School Proficiency Achievement Test (HSPA) and both Math and Verbal SAT scores. This study also examined such variables as: student mobility rate, District Factor Grouping, and percentage of students participating in the SAT. Current research has identified these factors as possible explanatory variables for student achievement.

Research Questions

This study attempted to answer the following research question:

How do students’ scores on standardized tests compare between New Jersey High Schools that operate state recognized service-learning programs and those that do not?

The research question was guided by the following sub questions:
1. How do New Jersey Leader and Community Service-Learning Schools HSFA scores compare to traditional high schools HSFA test scores in New Jersey?

2. How do New Jersey Leader and Community Service-Learning Schools SAT scores compare to traditional high schools SAT test scores in New Jersey?

3. Are there differences when comparing HSFA and SAT scores of New Jersey Leader Schools to Non-New Jersey Leader Schools?

Population

All of the 271 New Jersey public high schools that are included in the State Report Card 2002-2003 were considered initially for this study. In the fall of 2002, the Department of Education sent out a survey to every public, charter, and independent school in New Jersey. By January 2003, over 550 surveys were returned including 196 public high school respondents (n=196). These surveys were integrated into the study design when they were utilized to classify schools that have been identified by the Department of Education as fully implementing service-learning into the school culture. Of the 196 respondents to the survey, there were 27 CSL Schools and 149 schools that were not practicing service-learning. Also responding to the survey were the recipients of the New Jersey Service-Learning Leader School award (n=21). Recipients of the New Jersey Leader School award 2000-2003 by the New Jersey Department of Education were published on the Department website.

For the purposes of this study, the following sub-populations exist:

1. Public high schools with state-recognized service-learning programs
2. Public high schools without state-recognized service-learning programs
3. Privately run high schools, and adult high schools
Only groups 1 and 2 were utilized in this study.

Of total high school respondents practicing service-learning, 21 were New Jersey Leader Schools and 27 were Community Service-Learning Schools. From these samples the researcher removed vocational schools that did not list school test scores (in those instances the student population was tested in their sending district). Also while collecting the sample information, one of the charter schools included was closed. This left the lone charter school without an adequate match and consequently it was removed from the sample. The final sample used in the study includes 17 New Jersey Leader Schools and 22 Community Service-Learning Schools, which were matched to an additional 39 Traditional Schools that do not practice Community Service-Learning. A total number of 78 schools were part of the study.

When collecting the sample, each District Factor Grouping except DFG “1” was represented in the Leader School group. Of the 573 designated districts in New Jersey, there are 15 districts designated as “J”. The appendix sets forth the current distribution of DFG districts. The list of New Jersey Leader Schools (NJL) and Community Service-Learning Schools (CSL) used in this sample and their DFG designation are detailed in Tables 1 and 2.
<table>
<thead>
<tr>
<th>NLI School</th>
<th>DFG</th>
<th>NLI School</th>
<th>DFG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elizabeth High School</td>
<td>A</td>
<td>Middle Township School</td>
<td>B</td>
</tr>
<tr>
<td>Lieden High School</td>
<td>B</td>
<td>Lakewood High School</td>
<td>B</td>
</tr>
<tr>
<td>Kearny</td>
<td>B</td>
<td>Delaware Regional High School</td>
<td>CD</td>
</tr>
<tr>
<td>JFK High School</td>
<td>DE</td>
<td>Roselle Park High School</td>
<td>DE</td>
</tr>
<tr>
<td>Fair Lawn High School</td>
<td>FG</td>
<td>Middlesex Vo-Tech</td>
<td>FG</td>
</tr>
<tr>
<td>Vernon Township</td>
<td>FG</td>
<td>Canford High School</td>
<td>GH</td>
</tr>
<tr>
<td>Bemards High School</td>
<td>I</td>
<td>New Providence High School</td>
<td>I</td>
</tr>
<tr>
<td>Princeton High School</td>
<td>J</td>
<td>Whippany Park High School</td>
<td>I</td>
</tr>
<tr>
<td>Gloucester County Institute</td>
<td>V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSL School</td>
<td>DFG</td>
<td>CSL School</td>
<td>DFG</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----</td>
<td>--------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>Lincoln High School</td>
<td>A</td>
<td>Science High School</td>
<td>A</td>
</tr>
<tr>
<td>Bayonne High School</td>
<td>B</td>
<td>Manchester Twp. High School</td>
<td>B</td>
</tr>
<tr>
<td>Cliffside Park High</td>
<td>CD</td>
<td>Plumstead/New Egypt</td>
<td>CD</td>
</tr>
<tr>
<td>South Amboy Middle/High</td>
<td>DE</td>
<td>Woodbridge School District</td>
<td>DE</td>
</tr>
<tr>
<td>Colonia</td>
<td>DE</td>
<td>High Point Regional High</td>
<td>DE</td>
</tr>
<tr>
<td>North Arlington</td>
<td>DE</td>
<td>Hamilton Twp</td>
<td>FG</td>
</tr>
<tr>
<td>Rutherford High</td>
<td>FG</td>
<td>East Camden County Reg.</td>
<td>GH</td>
</tr>
<tr>
<td>Morristown High School</td>
<td>GH</td>
<td>Teaneck High School</td>
<td>GH</td>
</tr>
<tr>
<td>Cresskill Jr-St. High School</td>
<td>I</td>
<td>Glen Ridge High School</td>
<td>I</td>
</tr>
<tr>
<td>Hopewell High School</td>
<td>I</td>
<td>Northern Valley (Demarest)</td>
<td>I</td>
</tr>
<tr>
<td>Pascack Hills High</td>
<td>I</td>
<td>Charter-Tech</td>
<td>I</td>
</tr>
<tr>
<td>Salem County Ait High</td>
<td>V</td>
<td>Union County Magnet High</td>
<td>V</td>
</tr>
</tbody>
</table>

Given the limited number of schools available in this sample, the researcher recoded the DFG designations, consisting of eight groups (A, B, CD, DE, FG, GH, I, I) into three groups coded by numbers 1, 2, and 3 (Table 3).
Table 3

*Distribution of Redefined DFG Designations in Re-Coding for Study*

<table>
<thead>
<tr>
<th>Redefined DFG Code</th>
<th>DFG Designation</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>V, A, &amp; B</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>CD, DE, &amp; FG</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>I and GB</td>
<td>26</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>All</strong></td>
<td><strong>78</strong></td>
</tr>
</tbody>
</table>

The distribution by school type: Traditional, Community Service-Learning, NJ Leader, and Non-Leader are depicted below in Table 4.

Table 4

*Distribution of Schools by DFG*

<table>
<thead>
<tr>
<th>School Type</th>
<th>DFG Code</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>11</td>
<td>15</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Community Service-Learning (CSL)</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>New Jersey Leader Schools</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Non-Leader Schools</td>
<td>12</td>
<td>25</td>
<td>18</td>
<td></td>
</tr>
</tbody>
</table>

Data Collection

For this case study, data was collected through the state report card that is available through the New Jersey Department of Education website. These scores included SAT and HSPA scores from the spring of 2003 for each of the NJ Leader and...
the CSL schools and their DFG match. From the Department of Education website, data will also be gathered from HSPA results regarding the average scale score, the percent of students at or above state goal. From the state report card, information on student mobility rate, geographic location, and participation of students in testing was also gathered for the NJ Leader and CSL schools and the traditional school match. Further information is available via the College Board website on SAT scores with disaggregate data for gender, race, and score. This breakdown is available for all states.

Data Analysis

Analysis of this data involved a variety of descriptive and inferential statistics. Descriptive statistics were used to demonstrate the associations between variables in the study. Summary statistics of the data were created using Statistical Package for the Social Sciences (SPSS), a data management tool, and Excel software. The data was recorded on excel worksheets. The mean and standard deviations were calculated for each subgroup.

The inferential testing method used was the Analysis of Covariance (ANCOVA). ANCOVA is a control technique used in both causal-comparative studies in which already formed but not necessarily equal groups that are involved, and in experimental studies in which either existing groups or randomly formed groups are involved.

Analysis of covariance increases the power of a statistical test by reducing within-group variance. ANCOVA can reduce random sampling error by equating different groups and it increases the power of the significance as the power is directly related to the degree of randomization involved in formation of the groups.

Analysis of Covariance (ANCOVA) was utilized to analyze the main effect of the New Jersey Leader School and Community Service-Learning Schools and any interaction
between the school variable and DFG assignment while applying statistical controls in comparing the test score data. An Analysis of Covariance is control techniques used in both causal-comparative studies in which already formed but not necessarily equal groups are involved (Gay & Airasian, 2, p. 498). Use of ANCOVA allows for initial differences on a variable to be equated with respect to the control variable and then compared. By using ANCOVA the researcher attempted to reduce variation in posttest scores that is attributable to another, non-treatment, variable.

In this study, the primary independent variables are New Jersey Leader Schools and Community Service-Learning Schools. District Factor Grouping is also an independent variable tested for interaction. The covariates are student mobility rate, and the percentage of the student body participating in SAT testing. The dependent variable is student test scores for the verbal and math portions of the HSPE and SAT tests. The use of these statistical treatments will allow the researcher to test the null hypothesis presented in the study and provide answers to the research questions. Data gathered in this study will be statistically analyzed utilizing SPSS software and coded as indicated in Table 5.

The relevant statistical technique used to analyze each of the null hypotheses is stated below. The significance of each hypothesis was determined at the .05 probability level, which is the standard for social research.
<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>Data Source</th>
<th>Statistical Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Students in a New Jersey Leader or Community Service-Learning school will perform no differently than students in a traditional high school on the verbal portion of the HSPA test.</td>
<td>Mean scores</td>
<td>Descriptive statistics</td>
</tr>
<tr>
<td>2. Students in a New Jersey Leader or Community Service-Learning school will perform no differently than students in a traditional high school on the math portion of the HSPA test.</td>
<td>Mean scores</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>3. Students in a New Jersey Leader or Community Service-Learning school will perform no differently than students in a traditional high school on the verbal portion of the SAT test.</td>
<td>Mean scores</td>
<td>ANCOVA</td>
</tr>
</tbody>
</table>
4. Students in a New Jersey Leader or Community Service-Learning school will perform no differently than students in a traditional high school on the math portion of the SAT test.

5. Students in a New Jersey Leader or Community Service-Learning school will perform no differently than students in a traditional high school on the total scores of the SAT test.

6. Students in a New Jersey Service-Learning Leader school will perform no differently than students in a Non-Leader high school on the verbal portion of the HSPA test.

7. Students in a New Jersey Service-Learning Leader school will perform no differently than students in a Non-Leader high school on the math portion of the HSPA test.
3. Students in a New Jersey Service-Learning Leader school will perform no differently than students in a Non-Leader high school on the verbal portion of the SAT test.

9. Students in a New Jersey Service-Learning Leader school will perform no differently than students in a Non-Leader high school on the math portion of the SAT test.

10. Students in a New Jersey Service-Learning Leader school will perform no differently than students in a Non-Leader high school on total scores of the SAT test.

Matching of the Sample

For this study, each Leader and Community Service-Learning (CSL) school was matched with a similar school within the Leader/CSL school district. If a similar school was not available, then a randomly selected traditional school with the same DFG was utilized. Vocational, Alternative, and Traditional schools were matched with similarly themed schools.
The matching of the schools was performed by first utilizing the Department of Education Service-Learning Survey. Via the survey, the high school respondents were identified. The high school respondents were then assigned a group designating their status as a Leader School, CSL School or a non service-learning school. Using the State Report Card, the enrollment of each school was determined. Schools were then assigned a grouping number based on their DFG and size. Leader schools and CSL schools were then matched to non service-learning schools within the same district. If a match was not available within the district, the match was chosen within the county based on the school DFG assignment.

F-tests were employed to determine homogeneity of variance. An F-test is used to test if the standard deviations of two populations are equal (Witte & Witte, 2001, p.366). This test can be a two-tailed test or a one-tailed test. The two-tailed version tests against the alternative that the standard deviations are not equal. Independent sample t-tests were performed to determine whether the matches were equivalent by school demographics. A significant mean difference between matched schools for District Factor Grouping, student mobility rate, or percentage of participation in the SAT would denote an inequality in the school match and signify a non-homogeneous sample. The formula for the t-test is the ratio of the difference between two means and the standard error of the differences (Witte & Witte, 2001, p.294). The denominator, the standard error or standard deviation, is a measure for the variability of the school demographic variable. A positive t-value would indicate that the first mean in the t-test is larger than the second. A negative value indicates it is smaller.
Demographic statistics for Community Service-Learning School (CLS) and Traditional School are displayed in Table 6.

Table 6

Descriptive Statistics for Schools

<table>
<thead>
<tr>
<th>Variable</th>
<th>School Type</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility</td>
<td>CLS</td>
<td>39</td>
<td>10.15</td>
<td>8.12</td>
</tr>
<tr>
<td></td>
<td>Traditional</td>
<td>39</td>
<td>9.61</td>
<td>7.30</td>
</tr>
<tr>
<td>Percentage SAT</td>
<td>CLS</td>
<td>39</td>
<td>80.75</td>
<td>16.35</td>
</tr>
<tr>
<td></td>
<td>Traditional</td>
<td>39</td>
<td>78.51</td>
<td>18.57</td>
</tr>
<tr>
<td>District Factor</td>
<td>CLS</td>
<td>39</td>
<td>4.81</td>
<td>2.39</td>
</tr>
<tr>
<td></td>
<td>Traditional</td>
<td>39</td>
<td>5.05</td>
<td>2.46</td>
</tr>
</tbody>
</table>

The mean difference of CLS mobility and traditional Schools shows a t value of .312 (df = 79) p > .05, which indicates that there is not a significant difference between the mobility means. The mean of the percentage of students participating in the SAT for CLS Schools and Traditional Schools is also not significant with a t value of .569 (df = 77) p > .05, indicating that there is not a significant difference in the means. The District Factor Grouping also was not significant with a t value of -.499 (df = 80) p > .05. The bivariate t-test results were not statistically significant which confirms that the school matches are equivalent and will not modify the study outcomes (Table 7).
**Table 7**

**Independent t-Test Statistics for Schools**

<table>
<thead>
<tr>
<th>Variable</th>
<th>t</th>
<th>df</th>
<th>Sig (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility</td>
<td>.312</td>
<td>79</td>
<td>.756</td>
</tr>
<tr>
<td>Percentage SAT</td>
<td>.569</td>
<td>77</td>
<td>.571</td>
</tr>
<tr>
<td>District Factor</td>
<td>.499</td>
<td>80</td>
<td>.685</td>
</tr>
</tbody>
</table>

The mean difference of NJ Leader mobility and traditional Schools shows a t value of -.312 (df=79) p>.05, indicating that there is not a significant difference between the mobility means. The mean of the percentage of students participating in the SAT for NJ Leader Schools and Non-Leader Schools is also not significant with a t value of -.990 (df 77) p>.05, indicating that there is not a significant difference in the means. The District Factor Grouping also was not significant with a t value of .586 (df=80) p>.05. The bi-varient t-test results were not statistically significant (Table 8 and Table 9). This confirms that the school matches are equivalent and the will not modify the study outcomes.
Table 8

<table>
<thead>
<tr>
<th>Variable</th>
<th>School Type</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility</td>
<td>NJ Leader</td>
<td>17</td>
<td>8.83</td>
<td>6.51</td>
</tr>
<tr>
<td></td>
<td>Non-Leader</td>
<td>61</td>
<td>10.18</td>
<td>8.03</td>
</tr>
<tr>
<td>Percentage SAT</td>
<td>NJ Leader</td>
<td>17</td>
<td>75.94</td>
<td>20.65</td>
</tr>
<tr>
<td></td>
<td>Non-Leader</td>
<td>61</td>
<td>80.66</td>
<td>16.45</td>
</tr>
<tr>
<td>District Factor</td>
<td>NJ Leader</td>
<td>17</td>
<td>5.22</td>
<td>2.58</td>
</tr>
<tr>
<td>Grouping</td>
<td>Non-Leader</td>
<td>61</td>
<td>4.84</td>
<td>2.38</td>
</tr>
</tbody>
</table>

Table 9

Independent t-Test Statistics for School Demographic Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>t</th>
<th>df</th>
<th>Sig (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility</td>
<td>-.656</td>
<td>79</td>
<td>.514</td>
</tr>
<tr>
<td>Percentage SAT</td>
<td>-.996</td>
<td>77</td>
<td>.325</td>
</tr>
<tr>
<td>DFG</td>
<td>.586</td>
<td>80</td>
<td>.560</td>
</tr>
</tbody>
</table>
Scholastic Aptitude Test (SAT)

The Scholastic Aptitude Test (SAT) has been utilized for over 75 years as a reliable measure of college readiness (College Board, 2004). The SAT historically has been used to assist colleges and universities in making admission decisions.

The SAT test measures verbal and mathematical reasoning skills. The SAT math section covers arithmetic, algebra, geometry, probability, and requires students to apply mathematical concepts and use data literacy skills. SAT verbal sections emphasize reading. More than half of the verbal portion is devoted to passage reading questions. Vocabulary is also assessed in sentence completion and analogy questions (College Board, 2004).

The SAT has been used as a predictor of first-year College grades (College Board, 2004). The College Board website provides research showing that the correlation between SAT scores and freshman grades is +0.52. Table 10 provides test characteristics for the SAT.

Table 10

<table>
<thead>
<tr>
<th></th>
<th>Verbal</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliability coefficient</td>
<td>.91-.93</td>
<td>.91-.93</td>
</tr>
<tr>
<td>Standard error of measurement (SEM)</td>
<td>29-32</td>
<td>30-33</td>
</tr>
<tr>
<td>Standard error of difference (STD)</td>
<td>40-45</td>
<td>41-46</td>
</tr>
<tr>
<td>Difficulty (average percent correct)</td>
<td>.51-.59</td>
<td>.52-.60</td>
</tr>
</tbody>
</table>

*Note*: (College Board, 2004)
College Board (2004) reported that 75,915 students in New Jersey participated in the SAT test for school year 2002-2003 with 85% of the participating students having a B average (80) grade point average of higher. The state average was recorded as a score of 507 for the verbal portion and 519 for math.

The New Jersey Department of Education however, reported that 60,124 students took part in the test with 76% of the participating students having a B average or higher (Department of Education, 2004). The state average was reported to be a 500 for the verbal portion and 519 for math.

**New Jersey High School Proficiency Assessment (HSPA)**

In 1998, the New Jersey legislature passed legislation (N.J.S.A. 18A: 7C-6.2) that requires all students who graduate from a public high school in New Jersey to demonstrate mastery of skills (New Jersey Department of Education, 2004). To fulfill this requirement, the Department of Education administered the High School Proficiency Test (HST 11) from 1993 to 2001 to all New Jersey eleventh grade students. In 2000, the department replaced the HST 11 with the High School Proficiency Assessment (HSPA) for students who enter the eleventh grade on or before September 1, 2001.

The HSPA is a state test given primarily to students in the 11th grade to measure knowledge and skills that are identified in the New Jersey Core Curriculum Standards. The students generally take the test in March and receive test scores in Mathematics and Language Arts Literacy. All New Jersey students, even those with Limited English proficiency (LEP) and Special Education students, must prepare for the test. A small percentage is exempted through their Individualized Education Programs (IEP). The results of the HSPA also satisfy federal requirements under the No Child Left Behind Act.
(NCLB Act of 2001 PL n# 107-110). New Jersey reported that 88,320 students participated in testing for school year 2002-2003. The state average was 80.1 for the verbal portion and 55.9 for the math portion (Department of Education, 2003).

Summary

This chapter described the design of the study. The purpose of the study, research questions and the null hypothesis were reviewed. The data analysis that was performed was discussed. The threats to validity common to causal comparative studies were identified and addressed. In the chapters following, the collected data will be presented and analyzed. The paper will conclude with a study summary and recommendations.
CHAPTER IV
PRESENTATION AND ANALYSIS OF DATA

Summary of Research

The purpose of this study was to examine how service-learning affects academic achievement of New Jersey High School students’ tests scores on the High School Proficiency Achievement Test (HSPT), and the SAT. While service-learning is practiced in over 62% of schools in the United States, little research has been focused upon the impact of service-learning upon academics (Furco, 2002).

This study also examined school demographic variables such as student mobility rate and district factor grouping given that current research has identified these factors as possible explanatory variables for student achievement (Fisher et al., 1996; Tencks & Phillips, 1998; Skaudera & Sosa, 2001). A factor related to successful academic achievement is regular attendance, or stable mobility in the same school (Kenhow, 1996; Temple & Reynolds, 1999). Average achievement scores of schools with a high rate of student mobility are much lower than schools with more stable enrollments (Beck, Kratzer & Jaken, 1997; Mao et al., 1998; Neuman, 1987; Vail, 1996). Not only are higher rates of mobility related to lower achievement (e.g., Mao, Whitsett, & Mellor, 1998), mobile students exhibit more behavior problems, academic deficiencies and social inadequacies in comparison to stable students (Alexander, Dauber & Enzweiler, 1996; Brown & Orthner, 1990; Taggart, Scarrman & Eckerling, 1989).
With high stakes testing rising in popularity, the researcher has chosen to utilize standardized testing results as a measure of academic achievement. If the results of this study indicate that service-learning programming has a strong impact on academic achievement in districts may choose to adopt policies of implementing more service-learning programming.

The SAT test is the single most-used standardized test of its type in the country (College Board, 2002). The SAT I: Reasoning Test, a multiple-choice exam with verbal and mathematics components, is used in combination with high school grades to predict a student’s readiness for college. Mean SAT scores are tracked annually as indicators of the quality of education in school districts, states and nationwide (Bracey, 1996; Powell & Steelman, 1996).

Nevertheless, when comparing data from different districts, counties, and states it is also important to take the percentage of student population taking the SAT into account. This number represents the percentage of 10th and 11th grade students with high academic aspirations according to a report from The College Board on SAT test takers. In 2003, more students took the SAT than ever before. The 1.4 million SAT takers in 2002-2003 indicates the largest increase in the number of SAT takers in more than 15 years (College Board, 2004). According to College Board, Thirty-eight percent of those SAT takers are first-generation college-bound students, and Forty-two percent of 2003 college-bound seniors reported grade averages of A+, A, or A-. The average grade point average of SAT participants was 3.29. Thus the percentage of the school population that actively seeks college admission is not only an indicator of achievement but like student mobility rate, could have a potential impact on this study and needs to be considered.
This study will also control for socioeconomic status (SES) via school District Factor Grouping (DFG) assignment. Current literature has demonstrated that SES is a large factor on student achievement and on issues related to the achievement gap (Streichland & Alvermann, 2004). Achievement disparities between DFG and race/ethnic groups occur annually on assessments in language arts literacy (NJ Department of Education, 2003). In a study of 13,600 schools and 2,300 districts, it was found that the odds of getting high test scores are improved by high SES and made worse by low SES (Howley & Dickel, 2002).

In this study, the mean school scores of achievement tests from the New Jersey High School Proficiency Assessment (HSPA) and the SAT from the 2002-2003 school year reports were used as the variables for academic achievement. Three research questions are supported by three subsidiary questions and 10 null hypotheses will be analyzed and discussed in the present and following chapter.

Description of the Sample

Of the total population of 271 public and charter high schools in New Jersey, 196 high schools responded to the New Jersey Department of Education Survey. In the fall of 2002, the Department of Education sent out a survey to every public, charter, and independent school in New Jersey. By January 2003, over 550 surveys were returned, including 196 public high school respondents (n=196).

Of the pool of respondents, 96 schools were identified for this study. The initial 22 schools were chosen because of their identity as New Jersey Service Learning Leader Schools (NJ Leader). Another 27 schools were added to the pool at Community Service-Learning Schools (CSL Schools) because of their demonstrated commitment to service-
learning based on series of survey answers including: 50% of the student body involved in service-learning, the integration of service-learning in the school curriculum and culture, and the inclusion of essential elements such as reflection. Only schools that met the State Survey standards were considered Community Service-Learning schools (CSL).

The remaining schools that did not participate in service-learning school programs were placed in a pool to be matched to the 48 service-learning schools by Distract Factor Grouping and county. These schools are identified herein as Traditional Schools. Schools that did not provide a complete data set were removed from the study. The total number of schools used in the study was 78.

Given the limited number of schools available in this sample, the researcher re-coded the DFG designations, consisting of nine groups (V, A, B, CD, DE, FG, GH, HI, J) into three groups coded by numbers 1, 2, and 3 named “RedefineDFG” (ReDFG) (Table 11).

Table 11

<table>
<thead>
<tr>
<th>ReDFG Code</th>
<th>DFG Designation</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>V, A, &amp; B</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>CD, DE, &amp; FG</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>I and GH</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>All</td>
<td>78</td>
</tr>
</tbody>
</table>

Distribution of RedefinedDFG Designations in Re-Coding
Thus the variable RedefinedDFG consists of three sub-grouping of ReDFG designations with ReDFG 1 consisting of the designations for the poorest schools and ReDFG 3 consisting of designations of districts from wealthier socioeconomic areas.

Information regarding HSRA and SAT scores was then gathered for each school from the New Jersey State Report Card Database. In some instances, technical schools did not provide test data because their students were tested in the schools of their sending districts and thus these schools were removed from the sample. The researcher gathered achievement indicators including HSRA math and verbal proficiency scores, SAT math and verbal averages. In addition to test scores, the researcher collected information on student mobility rate, and the percentages of students who participated in testing from the School Report Card database. This information was gathered for the Service-Learning Leader, Community Service-Learning and Traditional Schools and entered into an excel spreadsheet.

The descriptive statistics in Table 12 consist of the minimum, maximum, and mean scores reported by schools included in the study. Also included for comparison are the New Jersey State School mean test scores for 2002-2003. The maximum test scores for schools included in this study are considerably higher than the state mean. The overall means for each achievement indicator of schools included in this study however are generally lower than the State mean. Of the variables representing schools included in the study, the mean HSRA Verbal scores is slightly lower (80.88 v. 80.1), as is the SAT Verbal (518 v. 493.4), and Total SAT Scores (1,128 v. 1,004.90). It should be noted that the HSRA Math means are slightly higher (66.62 v. 65.9), as are the SAT Math (511.76 v.
500). These statistics shall be relevant when discussing the application of the study results to educational policy in Chapter V.

**Table 12**

**Descriptive Statistics of Test Scores and State Mean Test Scores**

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>NJ State Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSPA Verbal</td>
<td>81</td>
<td>9.19</td>
<td>100.00</td>
<td>20.88</td>
<td>15.77</td>
<td>80.1</td>
</tr>
<tr>
<td>HSPA Math</td>
<td>81</td>
<td>3.00</td>
<td>100.00</td>
<td>66.62</td>
<td>22.38</td>
<td>65.9</td>
</tr>
<tr>
<td>SAT Verbal</td>
<td>80</td>
<td>348.00</td>
<td>630.00</td>
<td>493.14</td>
<td>54.85</td>
<td>518</td>
</tr>
<tr>
<td>SAT Math</td>
<td>80</td>
<td>373.00</td>
<td>645.00</td>
<td>511.76</td>
<td>60.51</td>
<td>500</td>
</tr>
<tr>
<td>SAT Total</td>
<td>80</td>
<td>721.00</td>
<td>1,273.00</td>
<td>1,004.96</td>
<td>113.88</td>
<td>1,128</td>
</tr>
</tbody>
</table>

**Analysis of Data**

The purpose of this chapter is to present the results of the analysis of the descriptive statistics using a series of ANCOVA models. In this study, the primary independent variables are levels of Service-Learning programming including the categories of Service-Learning Leader School, Community Service-Learning School, and Traditional school (Traditional Schools having no service-learning programming). Two statistical methods were used to analyze the data: *t*-tests, and Analysis of Covariance. The *t*-tests were performed to analyze the match of CSL and traditional school and homogeneity of variance. This process was used and discussed in detail in Chapter III.

Analysis of Covariance is used in many situations to determine whether there are differences between groups on the basis of one or more outcome variables or if a
continuous variable is a good predictor of one or more dependent variables. The univariate general linear model compares differences between group means and estimates the effect of covariates on a single dependent variable. Analysis of covariance is a popular procedure for removing extraneous error variance and increasing the power associated with the F test and for adjusting for the preexisting differences between groups (Conston & Colman, 2003). The null hypothesis is rejected if the observed F ration exceeds the critical value for F.

Analysis of Covariance (ANCOVA) was utilized to analyze the main effect of Service-Learning programming and any relationships between Service-Learning programming and District Factor Grouping (ReDFG) in terms of the dependent variable of achievement scores while applying statistical control to the variables of student mobility rate and percentages of students tested on the SAT test. The statistical tests were performed on the main variable of service-learning programming which includes: New Jersey Service-Learning Leader Schools (NJ Leader), Community Service-Learning Schools (CSL) and Traditional Schools.

ANCOVA was used as a tool to test the main effects of service-learning programming after controlling for identified variables of student mobility rate and percentage of students tested on the SAT. ANCOVA was also used to test whether the differences in the group means were significant.

Analysis of Covariance is an extension of ANOVA that allows for the exploration of differences between groups while statistically controlling for additional continuous variables. These additional variables (i.e., the covariates) are variables that may be influencing the dependent variable. ANCOVA regression procedures remove the
variation in the dependent variable that is due to the covariates, and, therefore, increases the sensitivity of the F-test.

The study utilized 10 null hypotheses to analyze the research data. Each hypothesis is listed separately followed by the results of the statistical treatment of the data. The rejection or acceptance of the hypothesis is based on the pre-determined significance level of \( p < .05 \).

**Research Questions and Hypotheses**

**Research Question 1:** How do New Jersey Leader and Community Service-Learning School students' HSFA test scores compare to Traditional New Jersey high school students' HSFA test scores?

**Null Hypothesis 1**

Null Hypothesis 1: Students in a New Jersey Leader or Community Service-Learning school will perform no differently than students in a Traditional School on the verbal portion of the HSFA test.

Descriptive statistics reveal the total number of schools tested was 77. Of the 77 tested schools, 39 were Community Service-Learning Schools and 38 were Traditional Schools. The HSFA Verbal Portion mean score of Community Service-Learning Schools was 84.71 (SD 11.22, N=39) compared to Traditional mean score 79.90 (SD 15.29, N=39) (Table 13).
Table 13

Group HSPA Verbal Scores

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community</td>
<td>39</td>
<td>84.7132</td>
<td>11.21693</td>
</tr>
<tr>
<td>Service-Learning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Community Service-Learning</td>
<td>38</td>
<td>79.007</td>
<td>15.19249</td>
</tr>
</tbody>
</table>

A 2 x 3 (CLS x RedDFG) between-groups ANCOVA was used to analyze these data in terms of the dependent variable of mean HSPA verbal scores. This analysis was performed using student mobility rate, and percentage of students tested on the SAT as statistical control. Significant effects were further analyzed using Least Square Difference (LSD) (p. >.05).

The results of the ANCOVA indicated significant covariates in student mobility rate F(1,69)=33.369, p<.000 and percentage of students participating in the SAT F(1,69)=6.129, p=.016 (p<.05). Because the covariates were significant, we knew that they were useful as statistical controls and the removing these covariate effects will allow further testing for the main and interaction effects (Table 14).

The main effect of Community Service-Learning was also significant F (1, 69) =7.582, p=.008. The main effect of Redefined DFQ however, was not significant with an F score of .905 (2, 69), p=. 409. The interaction effect between Community Service-Learning and RedefinedDFG did not yield a statistically significant values F (2, 69) =2.377, p=. 100.
Table 14

Analysis of Covariance for HSPA Verbal Scores

<table>
<thead>
<tr>
<th>Effect</th>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility</td>
<td>Between Subjects</td>
<td>1</td>
<td>2252.830</td>
<td>33.369</td>
<td>**.006</td>
</tr>
<tr>
<td>Percentage of Students</td>
<td>Between Subjects</td>
<td>1</td>
<td>413.792</td>
<td>6.129</td>
<td>*.016</td>
</tr>
<tr>
<td>Main Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community Service-Learning</td>
<td>Between Subjects</td>
<td>1</td>
<td>511.911</td>
<td>7.382</td>
<td>**.008</td>
</tr>
<tr>
<td>Redefined DFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>122.239</td>
<td>.905</td>
<td>.409</td>
</tr>
<tr>
<td>Interaction Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSL*ReDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>320.953</td>
<td>2.577</td>
<td>.100</td>
</tr>
</tbody>
</table>

Note: *$p<.05$. **$p<.01$.

These results indicate that the null hypothesis is rejected due to statistically significant mean differences between HSPA Verbal scores of Community Service-Learning Schools and Traditional Schools ($p=.008$) (Table 14). The Community Service-Learning Schools outscored Traditional Schools by an average of 6.69 points on the verbal portion of the HSPA test.

Null Hypothesis 2

Null Hypothesis 2: Students in a New Jersey Leadert or Community Service-Learning School will perform no differently than students in a Traditional School on the math portion of the HSPA test.
Table 15

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community Service-Learning</td>
<td>39</td>
<td>71.1711</td>
<td>18.25191</td>
</tr>
<tr>
<td>No Community Service-Learning</td>
<td>39</td>
<td>64.4846</td>
<td>22.57938</td>
</tr>
</tbody>
</table>

Of the 78 schools tested, there were 39 Community Service-Learning Schools and 39 Traditional Schools. The HSPA Math Portion mean score of Community Service-Learning Schools was 71.17 (SD 18.25, N=39) compared to the Traditional School mean score of 64.48 (SD 22.58, N=39) (Table 15).

The ANCOVA revealed significant (p<.05) covariate effects in Mobility F (1,69)=38.798, p=.000 and percentage of students participating in the SAT F (1,69)=4.567, p=.036. The main effect of Community Service-Learning programming was significant at F (1, 69) =4.695, p=.034. However, the main effect of RedefinedDFG was not significant at F (2, 69) =2.295, p=.108, nor was the interaction effect of CSL and RedefinedDFG [an F (2, 69) =2.281, p=.110 (Table 16).

The positive statistically significant main effect of CSL indicates that the null hypothesis is rejected. There is significant mean difference between HSPA Math scores of Community Service-Learning Schools and Traditional Schools (p<.034) (Table 15). In the analysis, the Community Service-Learning Schools outscores Traditional Schools by an average of 5.7 points on the math portion of the HSPA test.
Table 16

Analysis of Covariance for HSPA Math Scores

<table>
<thead>
<tr>
<th>Effect</th>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility</td>
<td>Between Subjects</td>
<td>1</td>
<td>5474.791</td>
<td>38.598</td>
<td>**.009</td>
</tr>
<tr>
<td>Percentage Of Students</td>
<td>Between Subjects</td>
<td>1</td>
<td>644.670</td>
<td>4.569</td>
<td>.016</td>
</tr>
<tr>
<td>Main Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community Service-Learning</td>
<td>Between Subjects</td>
<td>1</td>
<td>662.544</td>
<td>4.695</td>
<td>.034</td>
</tr>
<tr>
<td>Redefined DFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>647.618</td>
<td>2.295</td>
<td>.108</td>
</tr>
<tr>
<td>Interaction Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSL*ReDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>643.798</td>
<td>2.281</td>
<td>.110</td>
</tr>
</tbody>
</table>

Note. *p<.05. **p<.01.

Research Question 2: How do New Jersey Leader and Community Service-Learning School students’ SAT test scores compare to Traditional New Jersey high school students’ SAT scores? (see Table 17 and Table 18).

Null Hypothesis 3

Null Hypothesis 3: Students in a New Jersey Leader or Community Service-Learning school will perform no differently than students in a Traditional Schools on the verbal portion of the SAT test.
Table 17

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community Service-Learning</td>
<td>39</td>
<td>501.2564</td>
<td>51.70855</td>
</tr>
<tr>
<td>No Community Service-Learning</td>
<td>39</td>
<td>482.5641</td>
<td>53.48674</td>
</tr>
</tbody>
</table>

Table 18

Analysis of Covariance for SAT Verbal Scores

<table>
<thead>
<tr>
<th>Effect</th>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility</td>
<td>Between Subjects</td>
<td>1</td>
<td>33330.496</td>
<td>37.903</td>
<td>**.000</td>
</tr>
<tr>
<td>Main Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community Service-Learning</td>
<td>Between Subjects</td>
<td>1</td>
<td>6465.393</td>
<td>7.352</td>
<td>**.008</td>
</tr>
<tr>
<td>RedefinedDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>32906.294</td>
<td>18.710</td>
<td>**.000</td>
</tr>
<tr>
<td>Interaction Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSL x ReDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>6530.540</td>
<td>3.713</td>
<td>*.029</td>
</tr>
</tbody>
</table>

Note. *p<.05. **p<.01.

This ANCOVA revealed statistically significant results with all of the variables tested. The mobility covariate F (1, 70) = 37.903, p = .000 was significant. The mean score of CLS schools (M=501.26, SD 51.71, N=39) was higher than the mean score of the Traditional schools (M=482.56, SD 53.90, N= 39) and this difference was significant.
F (1, 70) = 7.352, p = .008 (Table 17). The main effect of Redefined DFG was also significant F (2, 70) = 18.710, p = .000 with the highest scores in the RedFG 1 (M = 524.097) compared with the RedFG 2 scores (M = 484.085) and the RedFG 3 scores (M = 466.794). This means the highest means were found from schools in the highest socioeconomic communities. A posteriori LSD tests showed that the mean for the RedFG 3 was significantly higher than the means for the other two recoded DFG variables (p < .05, two-tailed). This pair wise difference was significant (Table 19).

Finally, the interaction effect of CSL and RedFG was also significant with an F score of (2, 70) = 2.683, p = .029. Thus Verbal SAT scores tended to be higher for schools with CSL programming than Traditional schools. This difference however, was most noted in the lowest socioeconomic districts which were coded RedFG 1. The largest difference was discovered in RedFG 1 Schools, which showed a difference of 39.1 points (Table 20). A review of the graph of the interaction was revealed to be ordinal.

For CSL schools in RedFG 3, the highest socioeconomic district grouping, scores tended to be higher than those scores from RedFG 1 Traditional schools. On average CSL schools showed higher scores of nearly 20.9 points.

CSL RedFG 2 district schools however, dropped slightly in scores with Community Service-Learning schools being negatively impacted by -4.7 points (Table 21). A review of the graph showed both RedFG 1 and 2 to be disordinal (Figure 1).
<table>
<thead>
<tr>
<th>ReDFG</th>
<th>Mean</th>
<th>Std. Error</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>466.794</td>
<td>6.337</td>
<td>453.164</td>
<td>480.424</td>
</tr>
<tr>
<td>2.00</td>
<td>484.085</td>
<td>5.566</td>
<td>472.990</td>
<td>495.179</td>
</tr>
<tr>
<td>3.00</td>
<td>524.097</td>
<td>6.264</td>
<td>511.610</td>
<td>536.584</td>
</tr>
</tbody>
</table>

Table 19

*Estimated Marginal Means Statistics for SAT Verbal Test*

<table>
<thead>
<tr>
<th>CSL</th>
<th>ReDFG</th>
<th>Mean</th>
<th>Std. Error</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>woCSL</td>
<td>1.00</td>
<td>447.282</td>
<td>8.831</td>
<td>429.478</td>
<td>465.086</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>486.753</td>
<td>8.274</td>
<td>470.259</td>
<td>503.246</td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td>513.633</td>
<td>8.621</td>
<td>496.448</td>
<td>530.818</td>
</tr>
<tr>
<td>CSL</td>
<td>1.00</td>
<td>486.307</td>
<td>9.368</td>
<td>467.632</td>
<td>504.982</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>481.417</td>
<td>7.417</td>
<td>466.531</td>
<td>496.293</td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td>534.560</td>
<td>8.481</td>
<td>517.653</td>
<td>551.467</td>
</tr>
</tbody>
</table>

Table 20

*Estimated Marginal Means Statistics by School for SAT Verbal Test*
Table 21

**Estimated Marginal Means for ReDFG by School on SAT Verbal Test**

<table>
<thead>
<tr>
<th>ReDFG</th>
<th>ReDFG</th>
<th>Mean diff.</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>2.00</td>
<td>-17.291</td>
<td>9.017</td>
<td>.059</td>
<td>-.35265</td>
<td>.684</td>
</tr>
<tr>
<td>3.00</td>
<td></td>
<td>-57.302</td>
<td>10.030</td>
<td>.000</td>
<td>-.77296</td>
<td>-.37308</td>
</tr>
<tr>
<td>2.00</td>
<td>1.00</td>
<td>17.291</td>
<td>9.017</td>
<td>.059</td>
<td>-.684</td>
<td>35.265</td>
</tr>
<tr>
<td>3.00</td>
<td></td>
<td>-40.012</td>
<td>8.220</td>
<td>.000</td>
<td>-.56397</td>
<td>-23.626</td>
</tr>
<tr>
<td>3.00</td>
<td>1.00</td>
<td>57.302</td>
<td>10.030</td>
<td>.000</td>
<td>37.308</td>
<td>77.296</td>
</tr>
<tr>
<td>2.00</td>
<td></td>
<td>40.012</td>
<td>8.220</td>
<td>.000</td>
<td>23.626</td>
<td>56.397</td>
</tr>
</tbody>
</table>
Figure 1. Group Verbal SAT Scores
These results indicate that the null hypothesis is rejected due to statistically significant mean differences between SAT Verbal scores of Community Service-Learning Schools and Traditional Schools ($p=.010$). The Community Service-Learning Schools outscored Traditional Schools by an average of 17.69 points on the verbal portion of the SAT test.

**Null Hypothesis 4**

Null Hypothesis 4: Students in a New Jersey Leader or Community Service-Learning school will perform no differently than students in a Traditional Schools on the math portion of the SAT test.

Table 22

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community Service-Learning</td>
<td>39</td>
<td>518.8462</td>
<td>53.43148</td>
</tr>
<tr>
<td>No Community Service-Learning</td>
<td>39</td>
<td>502.7436</td>
<td>63.92953</td>
</tr>
</tbody>
</table>

The results of the ANCOVA revealed statistically significant results for the mobility covariates $F(1, 70)=37.903, p=.000$ and percentage of students participating in the SAT $F(1, 70)=4.350, p=.000$. The mean score of CLS schools ($M=518.85, SD=53.43, N=43$) was higher than the mean score of the Traditional schools ($M=502.74, SD$).
63.92, N=39) but this difference was not significant F (1, 70) =3.514, p=.065 (see Table 22 and Table 23).

Table 23

**Analysis of Covariance for SAT Math Scores**

<table>
<thead>
<tr>
<th>Effect</th>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Covariate Effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility</td>
<td>Between Subjects</td>
<td>1</td>
<td>25888.438</td>
<td>18.932</td>
<td>**.000</td>
</tr>
<tr>
<td>Percentage Of students</td>
<td>Between Subjects</td>
<td>1</td>
<td>22195.578</td>
<td>18.860</td>
<td>**.000</td>
</tr>
<tr>
<td><strong>Main Effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community Service-Learning</td>
<td>Between Subjects</td>
<td>1</td>
<td>4135.636</td>
<td>3.514</td>
<td>.065</td>
</tr>
<tr>
<td>RedefinedDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>16747.044</td>
<td>.115</td>
<td>* .02</td>
</tr>
<tr>
<td><strong>Interaction Effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSL*RedefinedDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>8046.930</td>
<td>3.420</td>
<td>* .038</td>
</tr>
</tbody>
</table>

*Note: *p<.05. **p<.01.

The main effect of RedefinedDFG was significant F (2, 70) =3.420 p=. 065. The highest mean was from RedefinedDFG, which represents schools from communities of the highest socioeconomic level. RedefinedDFG 3 (M=539.449) compared with the RedefinedDFG 2(M=509.770), and RedefinedDFG 1(M=483.056) (see Table 24). A posteriori LSD tests showed that the mean for the RedefinedDFG 3 was significantly higher than the means for the
other two recoded DFG variables (p < .05, two-tailed). The pair wise differences between all three ReDFG were significant.

Finally, the interaction effect of CSL and ReDFG was significant F (2,70)=3.420, p< .05. Thus Math SAT scores tended to be higher in schools with CSL programming than Traditional schools. This difference was most noted in the lowest socioeconomic district ReDFG 1. The largest difference was discovered in ReDFG 1 Schools, which showed a difference of 38.8 points (Table 26). A review of the line graph revealed that the interaction was ordinal, and further analysis indicated that the difference in the scores between the ReDFG 1 grouping rose when CSL programming was indicated (Figure 2). However there was no significant difference in CSL and Traditional school scores overall (Table 25).

For CSL schools in ReDFG 3, the highest socioeconomic district grouping, scores tended to be higher than those scores from ReDFG 3 Traditional schools. On average CSL schools from ReDFG 3 showed higher mean scores of 16.2 points.

CSL ReDFG 2 district schools however, dropped slightly in scores with Community Service-Learning schools being negatively impacted by 11.7 points (Table 19). A review of the graph showed both ReDFG 1 and 2 to be disordinal.
### Table 24

**Estimated Marginal Means Statistics for SAT Math Test**

<table>
<thead>
<tr>
<th>ReDFG</th>
<th>Mean</th>
<th>Std. Error</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>483.056</td>
<td>9.193</td>
<td>464.722</td>
<td>501.390</td>
</tr>
<tr>
<td>2.00</td>
<td>509.770</td>
<td>6.550</td>
<td>496.707</td>
<td>522.833</td>
</tr>
<tr>
<td>3.00</td>
<td>539.449</td>
<td>8.715</td>
<td>522.066</td>
<td>556.831</td>
</tr>
</tbody>
</table>

### Table 25

**Estimated Marginal Means Statistics by School for SAT Math Test**

<table>
<thead>
<tr>
<th>ReDFG</th>
<th>ReDFG</th>
<th>Mean Diff (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>2.00</td>
<td>-26.714</td>
<td>11.358</td>
<td>.021</td>
<td>-49.366</td>
<td>-4.062</td>
</tr>
<tr>
<td>3.00</td>
<td></td>
<td>-56.392</td>
<td>15.009</td>
<td>.000</td>
<td>-86.327</td>
<td>-26.458</td>
</tr>
<tr>
<td>2.00</td>
<td>1.00</td>
<td>26.714</td>
<td>11.358</td>
<td>.021</td>
<td>4.062</td>
<td>49.366</td>
</tr>
<tr>
<td>3.00</td>
<td></td>
<td>-29.679</td>
<td>10.900</td>
<td>.008</td>
<td>-51.418</td>
<td>-7.939</td>
</tr>
<tr>
<td>3.00</td>
<td>1.00</td>
<td>56.392</td>
<td>15.009</td>
<td>.000</td>
<td>26.458</td>
<td>86.327</td>
</tr>
<tr>
<td>2.00</td>
<td></td>
<td>29.679</td>
<td>10.900</td>
<td>.008</td>
<td>7.939</td>
<td>51.418</td>
</tr>
<tr>
<td>ReDFG</td>
<td>CSL</td>
<td>Mean</td>
<td>Std. Error</td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>-------</td>
<td>------------</td>
<td>-------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>noCSL</td>
<td>463.633</td>
<td>11.973</td>
<td>439.754</td>
<td>487.512</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CSL</td>
<td>502.479</td>
<td>11.335</td>
<td>479.872</td>
<td>525.086</td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td>noCSL</td>
<td>515.630</td>
<td>9.569</td>
<td>496.545</td>
<td>534.716</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CSL</td>
<td>503.910</td>
<td>8.902</td>
<td>486.155</td>
<td>521.664</td>
<td></td>
</tr>
<tr>
<td>3.00</td>
<td>noCSL</td>
<td>531.040</td>
<td>11.220</td>
<td>508.662</td>
<td>553.418</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CSL</td>
<td>547.857</td>
<td>10.812</td>
<td>526.293</td>
<td>569.422</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2. Group SAT Math Scores
These results indicate that the null hypothesis is accepted. A statistically significant mean difference was not found between SAT Math scores of Community Service-Learning Schools and Traditional Schools ($p = .073$) (Table 27). According to the descriptive statistics, the Traditional Schools outscored Community Service-Learning Schools by an average of 9.49 points on the Math portion of the SAT test, but this difference was not significant at the .05 level. The significant mean difference in Redefined DFG and the interaction between Redefined DFG and Community Service-Learning shows that the higher scores overall were held by the wealthiest school districts that had implemented service-learning. The largest difference in scores was yielded by the poorest districts, which also revealed higher scores by schools with CSL programming. The average income districts depicted a slight decrease with CSL programming when compared to same socioeconomic grouping of traditional non-CSL schools.

Null hypothesis 5

Null Hypothesis 5: Students in a New Jersey Leader or Community Service-Learning school will perform no differently than students in a Traditional Schools when comparing the total scores of the SAT test.

Table 27

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community</td>
<td>39</td>
<td>1020.103</td>
<td>104.06991</td>
</tr>
<tr>
<td>Service-Learning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Community Service-Learning</td>
<td>39</td>
<td>985.3077</td>
<td>115.59370</td>
</tr>
</tbody>
</table>
The ANCOVA performed illustrated that both covariates: student mobility rate $F(1, 69) = 28.738$, $p = .000$ and percentage of students participating in the SAT $F(1, 69) = 4.302$, $p = .042$ were significant ($p < .05$) (Table 28).

The SAT total mean score of CLS schools (Verbal and Math portions combined), $(M=1020.10, SD = 104.06, N=39)$ was higher than the total mean SAT scores of the Traditional Schools $(M=985.31, SD = 115.59, N=39)$ and this difference was significant at $F(1, 70) = 5.497$, $p = .022$.

The main effect of Redefined DFG was also significant $F(2, 70) = 7.272$, $p = .001$. The highest mean scores were found in the ReDFG 3 $(M=1060.037)$ compared with the ReDFG 2 $(M=1002.270)$ and the ReDFG 1 $(M=975.423)$. A posteriori LSD tests showed that the mean for ReDFG 3, the grouping of schools from the highest socioeconomic areas was significantly higher than the means for ReDFG 1 which represented the lowest and ReDFG 2 grouping of the average districts $(p < .05, 	ext{tow-tailed})$. This pair wise difference was significant. The differences between ReDFG 1 and ReDFG 2 however were not.
Table 28

**Analysis of Covariance for SAT Total Scores**

<table>
<thead>
<tr>
<th>Effect</th>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility</td>
<td>Between Subjects</td>
<td>1</td>
<td>106939.989</td>
<td>28.738</td>
<td>**.000</td>
</tr>
<tr>
<td>Percentage Of Students</td>
<td>Between Subjects</td>
<td>1</td>
<td>16009.768</td>
<td>4.302</td>
<td>*.042</td>
</tr>
<tr>
<td>Main Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community Service-Learning</td>
<td>Between Subjects</td>
<td>1</td>
<td>20456.529</td>
<td>5.497</td>
<td>**.022</td>
</tr>
<tr>
<td>RedefinedDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>54117.342</td>
<td>7.272</td>
<td>**.001</td>
</tr>
<tr>
<td>Interaction Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSL*ReDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>24730.239</td>
<td>3.323</td>
<td>*.042</td>
</tr>
</tbody>
</table>

*Note.* *p<.05. **p<.01.

The interaction effect of CSL and ReDFG was also significant with an F score of 3.323(2, 70), p=.042. This indicates that there are significant differences at the p<.05 level (Table 29). Thus a closer examination of CSL schools and total SAT by RedefinedDFG groupings needed to be completed.

For CSL schools in ReDFG 1, the lowest socioeconomic district grouping, scores tended to be higher than those scores from ReDFG 1 Traditional schools (Table 30). On average CSL schools showed higher scores of nearly 73.513 points, which was the largest increase. The line graph of ReDFG 1 was disordinal (Table 31). ReDFG 3 yielded the
highest scores for CSL programming with a mean of 1075.339. The difference between Traditional and CSL schools was 38.672 points. A review of the graph of the interaction was revealed to be ordinal (Figure 3).

CSL ReDFG 2 district schools however, dropped slightly in scores with Community Service-Learning schools being negatively impacted by -14.465 points (Table 30). A review of the graph showed ReDFG 2 to be disordinal.

The null hypothesis is rejected due to a statistically significant mean difference between the total SAT scores of Community Service-Learning Schools and Traditional Schools ($p = 0.05$). The Community Service-Learning Schools outscored Traditional Schools by an average of 33.80 points on the SAT when both the math and verbal portions of the SAT test. The significant interaction between CSL and ReDFG ($p = 0.044$) indicated that the ReDFG region that schools were assigned influenced CSL and total SAT scores.

Table 29

<table>
<thead>
<tr>
<th>ReDFG</th>
<th>Mean</th>
<th>Std. Error</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>975.423</td>
<td>17.304</td>
<td>940.910</td>
<td>1009.935</td>
</tr>
<tr>
<td>2.00</td>
<td>1002.270</td>
<td>15.323</td>
<td>971.708</td>
<td>1032.832</td>
</tr>
<tr>
<td>3.00</td>
<td>1060.037</td>
<td>18.353</td>
<td>1023.433</td>
<td>1096.641</td>
</tr>
</tbody>
</table>
Table 30

Estimated Marginal Means Statistics by School for SAT Total

<table>
<thead>
<tr>
<th>ReDFG</th>
<th>CSL</th>
<th>Mean</th>
<th>Std. Error</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>noCSL</td>
<td>920.549</td>
<td>21.290</td>
<td>878.087</td>
<td>963.019</td>
</tr>
<tr>
<td></td>
<td>CSL</td>
<td>994.062</td>
<td>20.155</td>
<td>953.864</td>
<td>1034.260</td>
</tr>
<tr>
<td>2.00</td>
<td>noCSL</td>
<td>1002.508</td>
<td>17.016</td>
<td>968.570</td>
<td>1036.445</td>
</tr>
<tr>
<td></td>
<td>CSL</td>
<td>988.043</td>
<td>15.830</td>
<td>956.471</td>
<td>1019.614</td>
</tr>
<tr>
<td>3.00</td>
<td>noCSL</td>
<td>1036.657</td>
<td>19.952</td>
<td>996.865</td>
<td>1076.450</td>
</tr>
<tr>
<td></td>
<td>CSL</td>
<td>1075.339</td>
<td>19.226</td>
<td>1036.993</td>
<td>1113.684</td>
</tr>
</tbody>
</table>

Table 31

Estimated Marginal Means Statistics for ReDFG by School for SAT Total

<table>
<thead>
<tr>
<th>ReDFG</th>
<th>ReDFG</th>
<th>Mean Difference</th>
<th>Std. Error</th>
<th>Sig</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>2.00</td>
<td>-26.847</td>
<td>23.038</td>
<td>.248</td>
<td>-72.796</td>
<td>19.101</td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td>-84.614</td>
<td>29.124</td>
<td>.005</td>
<td>-142.701</td>
<td>-26.528</td>
</tr>
<tr>
<td>2.00</td>
<td>1.00</td>
<td>26.847</td>
<td>23.038</td>
<td>.248</td>
<td>-19.101</td>
<td>72.796</td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td>-57.767</td>
<td>24.011</td>
<td>.019</td>
<td>-105.656</td>
<td>-9.878</td>
</tr>
<tr>
<td>3.00</td>
<td>1.00</td>
<td>84.614</td>
<td>29.124</td>
<td>.005</td>
<td>26.528</td>
<td>142.701</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>57.767</td>
<td>24.011</td>
<td>.019</td>
<td>9.878</td>
<td>105.656</td>
</tr>
</tbody>
</table>
Figure 3. Group Total SAT Scores
Synopsis of Research Questions 1 and 2

Both research question one and research question two focused on the comparison of public high schools in New Jersey that have implemented full scale Community Service-Learning programs as established by the survey results from the New Jersey Department of Education. Statistical analysis revealed that there were significant differences between Community Service-Learning (CSL) Schools and Traditional Schools when controlling for the variables of student mobility rate and percentage of student participation in SAT testing. The results revealed higher mean scores for CSL schools in HSPA Verbal, HSPA Math, SAT Verbal, and total SAT scores.

In interactions between the CSL programming variable and the ReDFG grouping variable revealed that the wealthier socioeconomic districts yielded the highest mean scores. Of those wealthy districts, the ones with CSL programming schools in these districts were the highest. The interactions between CSL and ReDFG also revealed that the lowest socioeconomic schools with CSL programming yielded the largest difference in scores. However schools in average socioeconomic districts tended to drop in mean scores when CSL programming was present.

Research question three is focused question specifically on New Jersey Service-Learning Leader Schools (NJ Leader). NJ Leader Schools, consists of a small number of nominated schools that are part of the national initiative that recognizes high schools and middle schools for their excellence in service-learning. New Jersey nominates twenty schools annually including both middle, and high school.

While New Jersey Service-Learning Leader schools can include Public, Catholic, Private, or Charters, this study has focused on public high schools only. Extensive
service-learning research has reported greater impacts at the secondary level. Furthermore, Catholic and Private schools generally have instilled a culture of service as part of their mission statement. Such schools would be problematic to include in this comparison.

Self-reporting problems were prevented in the NJ Leader sample as the schools were chosen through an intensive application and interview process by a panel from the New Jersey Department of Education. Initially, applications are reviewed at the state level. The state department then forwards their nominations to the Corporation for National and Community Service. A national peer review panel reviews the applications and makes recommendations to the chief executive officer of the Corporation for National and Community Service who announces the winners in April of each year (Corporation for National Service, 2003).

The third research question therefore, addresses more specifically the impact of service-learning programming on New Jersey public high schools by narrowing the sample to include only programs that have achieved excellence in the field of service-learning.

Research Question 3: Are there difference in test score means when comparing the HSPA and SAT scores of New Jersey Service-Learning Leader Schools and Non-Leader Schools?

Null Hypothesis 6

Null Hypothesis 6: Students in a New Jersey Service-Learning Leader school perform no differently than students in a Non-Leader high school on the verbal portion of the HSPA test.
### Table 32

**Group HSPA Verbal Scores by Leader Schools (LS)**

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NJ Leader School</td>
<td>17</td>
<td>86.7131</td>
<td>11.92532</td>
</tr>
<tr>
<td>Non NJ Leader</td>
<td>61</td>
<td>80.7131</td>
<td>13.87902</td>
</tr>
</tbody>
</table>

### Table 33

**Analysis of Covariance for HSPA Verbal Scores by Leader Schools**

<table>
<thead>
<tr>
<th>Effect</th>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility</td>
<td>Between Subjects</td>
<td>1</td>
<td>1920.668</td>
<td>27.284</td>
<td>**.000</td>
</tr>
<tr>
<td>Percentage</td>
<td>Between Subjects</td>
<td>1</td>
<td>772.932</td>
<td>10.960</td>
<td>**.001</td>
</tr>
<tr>
<td>Students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NJ Leader</td>
<td>Between Subjects</td>
<td>1</td>
<td>556.956</td>
<td>5.071</td>
<td>**.028</td>
</tr>
<tr>
<td>Schools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RedefinedDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>17.687</td>
<td>.251</td>
<td>.779</td>
</tr>
<tr>
<td>Interaction Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NJ Leader*ReDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>88.640</td>
<td>1.259</td>
<td>.290</td>
</tr>
</tbody>
</table>

*Note. *p<.05. **p<.01.*
The ANCOVA covariates; student mobility rate $F(1, 69) = 27.284$, $p = .000$ and percentage of students participating in the SAT were significant with $F(1, 69) = 10.980$, $p = .001$.

The descriptive statistics for the Verbal portion of the HSPA test for New Jersey Service-Learning Leader Schools was 86.71 ($SD = 11.93, N=17$) compared to Non New Jersey Service-Learning Leader Schools mean score 985.31 ($SD = 15.59, N=61$) (Table 25). The main effect of New Jersey Leader Schools was significant at $F(1, 69) = 5.071$, $p = .028$ meaning that the mean score differences were significant. The main effect of Redefined $DFG F(2, 69) = .251$, $p = .779$ was not significant, nor was the interaction effect of NJ Leader and Redefined $DFG F(2, 69) = 1.259$, $p = .290$.

These results indicate that the null hypothesis is rejected due to statistically significant mean differences between the New Jersey Service-Learning Leader Schools Verbal HSPA scores and the Verbal HSPA scores of Non-New Jersey Service-Learning Leader Schools ($p = .028$). The NJ Leader Schools outscored Non-NJ Leader Schools by an average of 6 points on the verbal portion of the HSPA test (Table 34).

Null Hypothesis 7

Null Hypothesis 7: Students in a New Jersey Service-Learning Leader school perform no differently than students in a Non-Leader high school on math portion of the HSPA test.
Table 34

*Group HSPA Math Scores by Leader Schools (LS)*

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NJ Leader school</td>
<td>17</td>
<td>74.4312</td>
<td>18.69191</td>
</tr>
<tr>
<td>Non- NJ Leader</td>
<td>61</td>
<td>66.0410</td>
<td>20.98680</td>
</tr>
</tbody>
</table>

Table 35

*Analysis of Covariance for HSPA Math Score by Leader Schools*

<table>
<thead>
<tr>
<th>Effect</th>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility</td>
<td>Between Subjects</td>
<td>1</td>
<td>4734.092</td>
<td>33.901</td>
<td>**.000</td>
</tr>
<tr>
<td>Percentage Students</td>
<td>Between Subjects</td>
<td>1</td>
<td>1237.844</td>
<td>8.864</td>
<td>**.004</td>
</tr>
<tr>
<td>Main Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NJ Leader Schools</td>
<td>Between Subjects</td>
<td>1</td>
<td>904.097</td>
<td>6.474</td>
<td>*.013</td>
</tr>
<tr>
<td>Redefined DFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>312.288</td>
<td>1.118</td>
<td>.333</td>
</tr>
<tr>
<td>Interaction Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NJ Leader*ReDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>328.739</td>
<td>1.177</td>
<td>.314</td>
</tr>
</tbody>
</table>

*Note.* *p<.05, **p<.01.
The Analysis of Covariance revealed both covariates: student mobility rate with an F (1,69) = 33.901, p<.000 and percentage of students participating in the SAT F(1,69) = 8.864 p=.004 to be significant (p<.000). Descriptive statistics regarding the Math portion of the HSPA test showed that the mean score for New Jersey Service-Learning Schools was 74.43 (SD 18.69, N=17) compared to Non New Jersey Service-Learning Schools mean score 66.04 (SD 20.99, N=61) (Table 37). This difference was significant at F (1, 79) = 6.474, p=.013.

However, both the main effect of RedefinedDfG F = (2, 69) = 1.118, p=.313 and the interaction effect of NJ Leader and ReDfG were not significant F (2, 69) = 1.177 p=.314. This means that there were not significant differences between the redefined DfG groupings as a main effect in mean test scores or in mean test scores when comparing NJ Leader schools. These results indicate that the null hypothesis is rejected due to statistically significant mean differences between the New Jersey Service-Learning Leader Schools Math HSPA mean scores and the Math HSPA mean scores of Non New Jersey Service-Learning Leader Schools (p=.013). The NJ Leader Schools outscored Non-NJ Leader Schools by an average of 8.39 points on the verbal portion of the HSPA test (Table 35).

**Null Hypothesis 8**

Null Hypothesis 8: Students in a New Jersey Leader school perform no differently than students in a Non-Leader high school on the verbal portions of the SAT test.
Table 36

*Group SAT Verbal Scores by Leader Schools (LS)*

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NJ Leader school</td>
<td>17</td>
<td>509.4706</td>
<td>53.04022</td>
</tr>
<tr>
<td>Non-JL Leader</td>
<td>61</td>
<td>487.0164</td>
<td>52.50603</td>
</tr>
</tbody>
</table>

Table 37

*Analysis of Covariance for SAT Verbal Scores by Leader Schools*

<table>
<thead>
<tr>
<th>Effect</th>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility</td>
<td>Between Subjects</td>
<td>1</td>
<td>25821.510</td>
<td>30.669</td>
<td>**.000</td>
</tr>
<tr>
<td>Percentage Students</td>
<td>Between Subjects</td>
<td>1</td>
<td>5980.235</td>
<td>7.103</td>
<td>**.010</td>
</tr>
<tr>
<td>Main Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NJ Leader Schools</td>
<td>Between Subjects</td>
<td>1</td>
<td>7514.478</td>
<td>8.925</td>
<td>**.004</td>
</tr>
<tr>
<td>RedefinedDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>7642.773</td>
<td>4.539</td>
<td>*.014</td>
</tr>
<tr>
<td>Interaction Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NJ Leader*ReDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>2776.633</td>
<td>1.649</td>
<td>.208</td>
</tr>
</tbody>
</table>

*Note.* *p<.05. **p<.01.
The ANCOVA revealed statistically significant results for the covariates student mobility rate $F(1, 69) = 10.669$, $p = .001$ and percentage of students participating in the SAT $F(1, 70) = 7.103$, $p = .010$ (Table 37). The mean score of NJ Leader schools ($M = 509.47$, $SD = 53.04$, $N = 17$) was higher than the mean score of the Non-NJ Leader schools ($M = 487.02$, $SD = 52.51$, $N = 61$) (Table 36) and this difference was significant $F(1, 70) = 8.925$, $p = .004$. The main effect of Redefined DFG was also significant $F(2, 70) = 4.539$, $p = .014$ with the highest scores in the highest socioeconomic district labeled RedFG 3 ($M = 520.486$) compared with the middle districts of RedFG 2 ($M = 489.300$) and the lowest socioeconomic districts of RedFG 1 ($M = 483.302$) (Table 38). A posteriori LSD tests showed that the mean for the RedFG 3 was significantly higher than the means for the other two recoded DFG variables ($p < .05$, tow-tailed). This pair wise difference was significant (Table 39).

Finally, the interaction effect of NJ Leader and RedFG was not significant with an $F(2, 70) = 1.649$, $p = .200$ value indicating that there is not a significant difference at the $p < .05$ level. Thus, there were not any significant differences between Redefined DFG groupings and CSL programming within the groups.
Table 38

*Estimated Marginal Means Statistics for SAT Verbal Test by Leader Schools (LS)*

<table>
<thead>
<tr>
<th>redf</th>
<th>Mean</th>
<th>Std. Error</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>483.302</td>
<td>8.171</td>
<td>467.005</td>
<td>499.599</td>
</tr>
<tr>
<td>2.00</td>
<td>489.300</td>
<td>7.236</td>
<td>474.866</td>
<td>503.731</td>
</tr>
<tr>
<td>3.00</td>
<td>520.486</td>
<td>8.666</td>
<td>503.202</td>
<td>537.771</td>
</tr>
</tbody>
</table>

Table 39

*Estimated Marginal Means by ReDFG for SAT Verbal Test by Leader Schools (LS)*

<table>
<thead>
<tr>
<th>ReDFG</th>
<th>Difference</th>
<th>Std. Error</th>
<th>Stg.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>-5.998</td>
<td>10.879</td>
<td>.583</td>
<td>-27.695</td>
<td>15.699</td>
</tr>
<tr>
<td>2.00</td>
<td>-37.184</td>
<td>13.753</td>
<td>.009</td>
<td>-64.613</td>
<td>-9.756</td>
</tr>
<tr>
<td>3.00</td>
<td>-31.186</td>
<td>11.338</td>
<td>.008</td>
<td>-53.800</td>
<td>-8.573</td>
</tr>
<tr>
<td>1.00</td>
<td>5.998</td>
<td>10.879</td>
<td>.583</td>
<td>-15.699</td>
<td>27.695</td>
</tr>
<tr>
<td>3.00</td>
<td>-37.184</td>
<td>13.753</td>
<td>.009</td>
<td>9.756</td>
<td>64.613</td>
</tr>
<tr>
<td>2.00</td>
<td>-31.186</td>
<td>11.338</td>
<td>.008</td>
<td>8.573</td>
<td>53.800</td>
</tr>
</tbody>
</table>

These results indicate that the null hypothesis is rejected due to statistically significant mean differences between the New Jersey Service-Learning Leader Schools.
Verbal SAT scores and the Verbal SAT scores of Non New Jersey Service-Learning Leader Schools ($p = .004$). The NJ Leader Schools outscored Non-NJ Leader Schools by an average of 22.46 points on the verbal portion of the SAT test.

**Null Hypothesis 9**

Null Hypothesis 9: Students in a New Jersey Leader school perform no differently than students in a Non-Leader high school on the math portion of the SAT test.

Table 40

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NJ Leader school</td>
<td>17</td>
<td>522.2541</td>
<td>55.52338</td>
</tr>
<tr>
<td>Non-NJ Leader</td>
<td>61</td>
<td>507.5902</td>
<td>60.09004</td>
</tr>
</tbody>
</table>
Table 41

Analysis of Covariance for Math Scores on the SAT by Leader Schools

<table>
<thead>
<tr>
<th>Effect</th>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility</td>
<td>Between Subjects</td>
<td>1</td>
<td>18611.405</td>
<td>15.250</td>
<td>**.090</td>
</tr>
<tr>
<td>Percentage Student</td>
<td>Between Subjects</td>
<td>1</td>
<td>9558.270</td>
<td>7.832</td>
<td>**.007</td>
</tr>
<tr>
<td>Main Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NJ Leader Schools</td>
<td>Between Subjects</td>
<td>1</td>
<td>4872.122</td>
<td>3.992</td>
<td>*.050</td>
</tr>
<tr>
<td>RedefinedDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>5020.982</td>
<td>4.114</td>
<td>*.020</td>
</tr>
<tr>
<td>Interaction Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NJ Leader*ReDFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>1431.227</td>
<td>1.173</td>
<td>.316</td>
</tr>
</tbody>
</table>

Note: *p<.05. **p<.01.

The ANCOVA revealed covariates student mobility rate $F(1, 70) = 15.250, p=.000$ and percentage of students participating in the SAT $F(1, 70) = 7.832, p=.007$ to be significant. The descriptive statistics for the SAT Math Portion mean score of New Jersey Service Learning Leader Schools was calculated to be 522.29 (SD 55.52, N=17) compared to Non-Leader Schools mean score of 507.59 (SD 60.09, N=66) and this difference was significant $F(1, 70) = 3.992, p= .050$ (Table 39, Table 40).

The main effect of ReDFG was $F(2, 70) = 4.114, p=.020$ with the highest scores in the ReDFG 3 (M=539.551) compared with the ReDFG 2(M=512.970) and the ReDFG
A posteriori LSD tests showed that the mean for the ReDFG 3 was significantly higher than the means for ReDFG 1 this pair wise difference was significant as was the difference between ReDFG 3 and ReDFG2 (p<.05, tow-tailed). There was not a significant difference between ReDFG 2 and ReDFG 1 (Table 41).

Finally, the interaction effect of NJ leader and ReDFG was not significant with an F (2, 70) = 1.173, p=.316 indicating that there is significant difference between the variables at the p<.05 level.

Table 42

<table>
<thead>
<tr>
<th>ReDFG</th>
<th>Mean</th>
<th>Std. Error</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>492.121</td>
<td>9.838</td>
<td>472.500</td>
<td>511.742</td>
</tr>
<tr>
<td>2.00</td>
<td>512.970</td>
<td>8.712</td>
<td>495.595</td>
<td>536.345</td>
</tr>
<tr>
<td>3.00</td>
<td>539.551</td>
<td>10.434</td>
<td>518.741</td>
<td>560.361</td>
</tr>
</tbody>
</table>

Estimated Marginal Mean Statistics for SAT Math Test by Leader Schools
### Table 43

**Estimated Marginal Means Statistics by ReDFG on SAT Math Test by Leader Schools**

<table>
<thead>
<tr>
<th>ReDFG</th>
<th>Mean</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>-20.849</td>
<td>13.098</td>
<td>.116</td>
<td>-46.972</td>
<td>5.273</td>
</tr>
<tr>
<td>3.00</td>
<td>-47.430</td>
<td>16.558</td>
<td>.006</td>
<td>-80.453</td>
<td>-14.407</td>
</tr>
<tr>
<td>2.00</td>
<td>20.849</td>
<td>13.098</td>
<td>.116</td>
<td>-5.275</td>
<td>46.972</td>
</tr>
<tr>
<td>3.00</td>
<td>-2.583</td>
<td>13.651</td>
<td>.056</td>
<td>-53.806</td>
<td>.645</td>
</tr>
<tr>
<td>3.00</td>
<td>47.430</td>
<td>16.558</td>
<td>.006</td>
<td>14.407</td>
<td>80.453</td>
</tr>
<tr>
<td>2.00</td>
<td>26.580</td>
<td>13.651</td>
<td>.056</td>
<td>-645</td>
<td>53.866</td>
</tr>
</tbody>
</table>

These results indicate that the null hypothesis is not accepted. There is a statistically significant mean difference between the New Jersey Service-Learning Leader Schools Math SAT scores and the Math SAT scores of Non-New Jersey Service-Learning Leader Schools ($p < .05$). The NJ Leader Schools outscored Non-NJ Leader Schools by an average of 14.70 points on the Math portion of the SAT test (Table 42).

**Null Hypothesis 10**

Null Hypothesis 10: Students in New Jersey Leader schools will perform no differently than students in a Non-Leader high school when comparing total scores of the SAT test.
Table 44

Total Group SAT Scores by Leader Schools

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NJ Leader school</td>
<td>17</td>
<td>1031.765</td>
<td>107.36068</td>
</tr>
<tr>
<td>Non- NJ Leader</td>
<td>61</td>
<td>994.6066</td>
<td>111.06099</td>
</tr>
</tbody>
</table>

Table 45

Analysis of Covariance for SAT Total Scores by Leader Schools (LS)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility</td>
<td>Between Subjects</td>
<td>1</td>
<td>88276.938</td>
<td>23.379</td>
<td>** .000</td>
</tr>
<tr>
<td>Percentage Of Students</td>
<td>Between Subjects</td>
<td>1</td>
<td>30659.446</td>
<td>8.120</td>
<td>** .006</td>
</tr>
<tr>
<td>Main Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NJ Leader Schools</td>
<td>Between Subjects</td>
<td>1</td>
<td>24488.080</td>
<td>6.485</td>
<td>.013</td>
</tr>
<tr>
<td>Redefined DFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>33793.670</td>
<td>4.475</td>
<td>.015</td>
</tr>
<tr>
<td>Interaction Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NJ Leader*RedFG</td>
<td>Between Subjects</td>
<td>2</td>
<td>11110.485</td>
<td>1.471</td>
<td>.237</td>
</tr>
</tbody>
</table>

Note. *p<.05. **p<.01.
The ANCOVA revealed that both covariates: student mobility rate $F (1, 70) = 23.379, p = .006$ and percentage of students participating in the SAT $F (1, 70) = 8.120, p = .006$ were reported to be significant. The descriptive statistics for total group SAT scores revealed a mean score of 1031.77 (SD 107.36, $N = 17$) for New Jersey Service-Learning Leader Schools. The mean score for the Non-Leader Schools was 994.61 (SD 111.06, $N = 6$) (Table 46). The main effect of New Jersey Leader Schools was significant at $F (1, 70) = 6.485, p = .013$ meaning the mean score difference was also significant. The main effect of Redefined RedFG was significant with an $F (2, 70) = 4.475, p = .015$ (Table 47). The highest mean scores were found in the RedFG 3 ($M = 1060.037$) compared with the RedFG 2 ($M = 1002.270$) and the RedFG 1 ($M = 975.423$) (Table 48). A posteriori LSD tests showed that the mean for the RedFG 3 was significantly higher than the means for the other two recoded DFG variables ($p < .05$, two-tailed). This pair wise difference was significant.

The interaction effect of NJ Leader and RedFG was not significant with an $F (2,70)=1.471, p=.237$.

Table 46

*Estimated Marginal Means Statistics for Total Scores on the SAT* by Leader Schools

<table>
<thead>
<tr>
<th>RedFG</th>
<th>Mean</th>
<th>Std. Error</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>975.423</td>
<td>17.304</td>
<td>940.910</td>
<td>1009.935</td>
</tr>
<tr>
<td>2.00</td>
<td>1062.270</td>
<td>15.223</td>
<td>971.708</td>
<td>1032.832</td>
</tr>
<tr>
<td>3.00</td>
<td>1060.037</td>
<td>18.353</td>
<td>1023.433</td>
<td>1096.641</td>
</tr>
</tbody>
</table>
Table 47

<table>
<thead>
<tr>
<th>ReDFG</th>
<th>ReDFG</th>
<th>Mean</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>2.00</td>
<td>-26.847</td>
<td>23.038</td>
<td>.248</td>
<td>-72.796</td>
<td>19.101</td>
</tr>
<tr>
<td>3.00</td>
<td></td>
<td>-84.614</td>
<td>29.124</td>
<td>.005</td>
<td>-142.701</td>
<td>-26.528</td>
</tr>
<tr>
<td>2.00</td>
<td>1.00</td>
<td>26.847</td>
<td>23.038</td>
<td>.248</td>
<td>-19.101</td>
<td>72.796</td>
</tr>
<tr>
<td>3.00</td>
<td>1.00</td>
<td>-57.767</td>
<td>24.011</td>
<td>.019</td>
<td>-105.656</td>
<td>-9.878</td>
</tr>
<tr>
<td>3.00</td>
<td>2.00</td>
<td>84.614</td>
<td>29.124</td>
<td>.005</td>
<td>26.528</td>
<td>142.701</td>
</tr>
<tr>
<td>2.00</td>
<td></td>
<td>57.767</td>
<td>24.011</td>
<td>.019</td>
<td>9.878</td>
<td>105.656</td>
</tr>
</tbody>
</table>

The significant effect of New Jersey Leader School programming (.013) indicates that the null hypothesis is rejected. The NJ Leader Schools outscored Non-NJ Leader Schools by an average of 37.16 points on the total score of the SAT test.

Summary

In summary, Analysis of Covariance tests demonstrated significant variance in Community Service-Learning Schools mean test scores when controlling for the school demographic variables of student mobility rate and percentage of students tested on the SAT. A statistically significant difference in mean test scores was found for CSL. Schools in the verbal and math portion of the HSPA, the verbal portion of the SAT, and total SAT scores when controlling for mobility, and percentage of students participating.
In all three instances, the Community Service-Learning Schools mean scores were higher than those of Traditional Schools.

In two cases, there was a significant interaction between RedFlag groupings and the CSL School variable. Both the Verbal and the Math portion of the SAT scores showed interactions that were statistically significant. These interactions revealed that both RedFlag 1 and RedFlag 3 groupings showed a higher mean difference between CSL and Traditional Schools. For both RedFlag Code 1 and RedFlag Code 3, the CSL School mean was higher than the Traditional Schools. The highest socioeconomic grouping labeled RedFlag 3 revealed the highest means and the lowest socioeconomic grouping labeled RedFlag 1 revealed the highest mean differences between CSL and traditional schools. In both cases, the middle socioeconomic grouping labeled RedFlag 2 showed a small but higher mean score for Traditional Schools when compared to schools with CSL programming.

When the same statistical tests were performed on mean test scores of New Jersey Service-Learning Leader Schools (NJ Leader) and Non-Leader Schools, a statistically significant difference was found in the verbal and math portion of the HSPE, the verbal and math portion of the SAT, and total SAT scores. In all instances the NJ Leader Schools outscored the Non-Leader Schools. Table 48 summarizes the outcomes of all the variables tested.

Nine of the ten Null Hypotheses were rejected as they relate to the three subsidiary questions. These Hypotheses were rejected based on the data analysis conducted on each. The implications of these results as they relate to the initial research
question (How does service-learning impact academic achievement?) and three subsidiary questions will be discussed further in Chapter V.

Table 48

Summary of findings

<table>
<thead>
<tr>
<th>Groups</th>
<th>HSPA</th>
<th></th>
<th>SAT</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Verbal</td>
<td>Math</td>
<td>Verbal</td>
<td>Math</td>
</tr>
<tr>
<td>CSL</td>
<td>.008</td>
<td>.034</td>
<td>.008</td>
<td>.065</td>
</tr>
<tr>
<td>NJ Leader</td>
<td>.028</td>
<td>.013</td>
<td>.004</td>
<td>.050</td>
</tr>
</tbody>
</table>

Note. *p<.05  **p<.01.
CHAPTER V
SUMMARY AND RECOMMENDATIONS

Conclusion
With the 2004 enactment of the No Child Left Behind (NCLB) Legislation, the federal government has trained a spotlight on the national problems of education funding inequities. As schools continue to fail to make the Adequate Yearly Progress Benchmarks defined by the NCLB bill, local reporters will highlight education deficiencies and schools will be forced to rely further on standardized testing to prove improvement. The current NCLB legislation has also focused educators and community members on perceiving standardized tests scores as a superior measurement of student aptitude and achievement. Now more than ever, school standing is utmost importance as federal funding and taxpayer dollars are in jeopardy over annual school test scores. Because of this, any educational method employed currently by teachers will continually be assessed on what impact it may have on the curriculum standards and standardized scores. Additionally, any methodology that can assist with student aptitude and achievement will be welcomed.

As long as such inequities exist and are publicly discussed, there will continue to be a search for a better methodology. While the United States continues to fall behind internationally in subjects such as math and science and greater disparities are noted
between urban and suburban students, it is important more than ever to find remedies in new education methodologies.

Service-learning as an educational method is still a developing pedagogy. Nationally, only a limited number of studies have focused on service-learning's impact on student academics at the high school level. In New Jersey, no statewide studies focusing on service-learning have been completed or published. While service-learning has become well known as a tool to foster civic participation and character education, the impact it may have on academic achievement and aptitude is widely unknown.

The purpose of this study was to examine how service-learning programming affects academic achievement of New Jersey High School students' tests scores on the High School Proficiency Achievement Test (HSPA), and SAT math and verbal scores. This study also examined the impact of such variables as: student mobility rate, geographic location, District Factor Grouping, and percentage of students participating in SAT testing.

This researcher believed that if results of this study indicate positive impacts on student aptitude and achievement that such findings could in fact be used to transform policy and practice in the State of New Jersey. With assistance and advocacy from the state government, districts may decide to adopt service-learning programming to increase learning, as well as raising test scores.

Summary of Research Design
This study attempted to determine the impact of service-learning programming in New Jersey public high schools on student academic achievement. The study employed both national and state resources by utilizing the New Jersey Service-Learning Leader
school nominations that were chosen by an intensive application and interview process by a panel from the New Jersey Department of Education. An additional number of schools were also utilized via the school year 2002-2003 Department of Education Service-Learning Survey that included 196 public and charter high school respondents. These particular surveys were integrated into the study design to include schools that have been identified by the Department of Education as fully implanting service-learning into the school culture (CSL Schools) and schools that were not practicing service-learning.

The total number of schools were identified from the pool respondents for this study was 78. The initial 17 schools were chosen because of their identity as New Jersey Service Learning Leader Schools (NJ Leader). Another 22 schools were added to the pool as Community Service-Learning Schools (CSL) because of their demonstrated commitment to including Community Service-Learning activities as part of the curriculum. This information was collected from the Department of Education Service-Learning Survey. From this survey District Factor Grouping and geographic location matched an additional 39 schools to the original sample. The additional 39 schools were drawn from a pool of survey respondents that indicated that they did not practice service-learning. These schools were referred to as “Traditional Schools.”

Test score data was gathered from the State Report Card Database. In addition to test scores, the researcher gathered information on student mobility rate, and the percentages of students who participated in SAT testing from the School Report Card database. This information was gathered for the Service-Learning Leader, Community Service-Learning and Traditional Schools.
This study utilized 10 null hypotheses to analyze the research data. Three statistical methods were used to analyze the data: t-tests, Analyses of Variance, and Analysis of Covariance. The t-tests were performed to determine homogeneity of variance. In this study, independent sample t-tests were performed to determine whether the matches were equivalent by school demographics. A significant mean difference between matched schools for District Factor Grouping, student mobility rate, or percentage of participation in the SAT would denote an inequality in the school match and signify a non-homogenous sample. Significant means were not found for mobility rate or percentage of participation in the SAT.

An Analyses of Co-Variance (ANCOVA) was used to analyze the main effect of Service-Learning programming and any relationships between Service-Learning programming and achievement scores while applying statistical control to the variables of District Factor Grouping (DFG), class size, mobility, and percentages of students tested on the SAT test. The dependent variable in the study was student achievement. The 2002-2003 school year mean scores of the verbal and math portions of the SAT and New Jersey HSPE test were collected to be used as measures of student achievement for 2003.

Review of Findings and Interpretations

The research question governing the study was: How does students' scores on standardized tests compare between New Jersey high Schools that operate state-recognized Service-learning Programs and those that do not? The research question was guided by three subsidiary questions, which are listed below and followed by the study findings. None of the ten Null Hypotheses were rejected as they relate to the three
subsidiary questions. These Hypotheses were rejected based on the data analysis conducted on each.

1. How do NJ Leader and CSL School students' HSPA test scores compare to Traditional New Jersey high school students' HSPA scores?

The Analysis of Covariance tests demonstrated significant variance in Community Service-Learning Schools mean test scores when controlling for the school demographic variables of student mobility rate and percentage of students tested on the SAT. A statistically significant difference in mean test scores was found for CSL Schools in the verbal and math portion of the HSPA when controlling for mobility, and percentage of students participating. In both instances, the Community Service-Learning Schools mean scores were higher than those of Traditional Schools with 6.69 points on the verbal portion and 5.7 on the math.

2. How do NJ Leader and CSL School students' SAT test scores compare to Traditional New Jersey high school students' SAT test scores?

A statistically significant difference in mean test scores was found for CSL Schools in the verbal portion of the SAT and total SAT scores when controlling for mobility, and percentage of students participating. In the verbal portion, CSL schools on average scored 7.69 mean points higher than traditional schools. In total scores, the mean difference between CSL schools and traditional schools was 33.8 points.

The ANCOVA results for all of the SAT tests revealed that there was a significant interaction between the socioeconomic groupings of the school and the Community Service-Learning School variable. In all of the tests for the SAT, results illustrated interactions that were statistically significant. These interactions revealed that ReDFG 1,
which represented the lowest socioeconomic grouping, and RedFG 3, which represented
the highest socioeconomic grouping, showed a higher mean difference between
Community Service-Learning and Traditional Schools.

In the verbal portion, the results showed that the highest socioeconomic grouping
labeled RedFG 3 had the highest means overall group with a 20.9 points mean
difference and the lowest socioeconomic grouping labeled RedFG1 revealed the highest
mean differences between CSL and traditional schools scores with an average 39.1 points
mean difference.

The testing of the math portion however, revealed although the mean score of CSL
schools was higher than the mean score of the traditional schools, the difference was not
significant (p=.065). The interaction effect was significant with Math scores tending to be
higher in schools with CSL programming. In the lowest socioeconomic districts the
difference was a mean total of 38.8 points. The highest socioeconomic difference yielded
a mean difference of 16.2. Again the middle districts in the RedFG 2 group showed a
difference of 11.7 points with the non CSL schools having higher scores. Again, the
mean difference overall between CSL and Traditional Schools was not significant and
null hypothesis four was not rejected.

The total SAT scores test showed that RedFG 1 scores were higher in CSL schools
with a mean of 73.5 points. RedFG 3 yielded a mean difference of 38.7. The middle
socioeconomic schools showed a 14.5 points difference favoring the non CSL schools.

3. Are there differences when comparing HSPA and SAT scores of New Jersey
Leader Schools to Non-New Jersey Leader Schools?
The third research question addressed the impact of service-learning programming on New Jersey public high schools more specifically by narrowing the sample to include only programs that have achieved excellence in the field of service-learning. The Analysis of Covariance tests demonstrated significant variance in New Jersey Leader schools mean test scores when controlling for the school demographic variables of student mobility rate and percentage of students tested on both the HSPA and SAT.

When testing New Jersey Service-Learning Leader Schools (NJ Leader) and Non-Leader Schools, a statistically significant difference was found in the verbal and math portions of the HSPA. The mean differences were higher for the NJ Leader Schools with a score of 6 points on the verbal and 8.39 on the math section. A significantly significant difference was also found when examining the verbal and math portions of the SAT, and total SAT scores. In all instances the NJ Leader Schools outscored the Non-Leader Schools. The mean score differences were 22.5 for the verbal section, 14.7 points per the math portion and 37.2 points on total SAT scores.

Implications

The purpose of this study was to examine the impact of service-learning programming on New Jersey public high schools. The research questions were designed to guide the study to a series of conclusions based on service-learning programming. A review of related literature revealed that few studies had focused on the academic impact of service-learning in K-12 schools and that no studies had been conducted in New Jersey. The outcomes will be significant therefore not only to service-learning practitioners and all educators in New Jersey, but to educators nationally that are searching for pedagogical tools that will increase student achievement and aptitude.
The findings of this study are in concert with those of Dewey’s theories of authentic learning (1933; Elyer & Giles, 1999), Piaget’s reflective thinking and cognitive development (1952), Freire’s authentic learning (1970), Kolb’s perceiving and processing premises (1984), Lewin’s collective and reflective inquiry hypothesis (Lewin, 1951) as well as the personal development and empowerment beliefs underscored by contemporary researchers Billig & Putco (2000a; 2000b; Billig 2000; Billig 2002; Billig 2004).

Furthermore, these findings are consistent with the research compiled by Billig (2002, 2004), which demonstrates that service-learning will assist with positive effect on the personal development of students’ interpersonal development as well as acquisition of academic skills and knowledge.

This study also revealed interactions within the RedFG groupings and service-learning in some cases. The three interactions revealed that lower DFG socioeconomic districts that had adopted service-learning programs yielded the greatest difference in scores. These findings tie into the most recent report issued by the National Youth Leadership Council, Community-Service and Service-Learning in U.S. Public Schools (www.NYLC.org, 2004).

It was found in a survey of 1,759 principals (91-percent response rate) that although schools serving low-income students are less likely to engage students in service-learning, there is evidence that those that do tend to see a greater impact on academic outcomes. It was reported that principals from high poverty level districts found the impact of service-learning very important for improving core academics as well.
as a positive impact on achievement (63% and 42%). This finding is likely related to the amount of supports and policies that have been implemented.

In November, Assemblywoman Bonnie Watson Coleman introduced new legislation based on the original bill sponsored by Acting Governor Richard Codey. The bill establishes a three-year community service program for juniors in high school beginning in the 2005-2006 school year. The Bill was passed by the Senate (35-0) in December 2004 and is in its 2nd Reading of Assembly Committee. If enacted, this bill will establishes a three-year community service pilot program under which high school juniors in 30 schools selected by the Commissioner of Education will be required to complete a minimum of 15 hours of community service.

Although the reasoning behind this bill is admirable, community service is not service-learning. While there is definite value in introducing youth to the concepts of service to the community and civic participation, the actual cognitive connections are missing. The bill requires that the student complete the community service requirement during evenings, weekends, or summers, but not during school hours. Without a link to curriculum and the essential element of reflection, “True learning” as Dewey (1938) defined, as authentic, active, and social, will not take place. It is the researcher’s opinion that the results of this study should be considered when forming recommended policies on education and that service-learning rather than community service is adopted.

Much research has been published on the importance of policies and support in implementing new programs. An important aspect of the education system is the policy context within which schools operate. Policies establish the principles, values, and ideals, which influence a school’s educators. Because state-level policymakers are interested
improving student aptitude and achievement, the results of this study may assist with creating new policies that adopt service-learning programming rather than community service.

In the intervening time, the researcher suggests continued studies on service-learning as well as increased record keeping from participating school districts.

Recommendations for Future Research

Practitioners from the field have reported that much of the existing research concerning service-learning is self-reported from directors seeking grant evaluations. It is not always clear whether such data can be validated (Billig, 2004; Furco, 2002; Hecht, 2003; Howard, 2003; Melchior & Bals, 2003; Shomer, 2003). Service-learning researchers have recently called for increased rigorous and replicable research in an effort to create a strong platform to advocate for more service-learning programs (Billig, 2004; Furco, 2002; Hecht, 2003). While it is firmly believed that service-learning programs impact concepts of self-value, social justice, and community awareness (Billig & Conrad, 1997; Melchior, 1999; Morgan, 1998), measurable improvements in student achievement is far more important to administrators and school board members who must allocate spending based on the value of each program.

This study focused only on public and charter high schools. These schools are regulated by the State and therefore have a similar makeup of curriculum and student bodies. This study focused on high school statistics only. Melchoir (1999) and Morgan (2000) have found service-learning to have a lasting impact in high school students and not at the middle school level. The data collection was confined to New Jersey Public High School students that have participated in the HSPA and pre-determined SAT scores.
Although this study attempted to control for socio-economic factors through District Factor Grouping, variables such as gender, outside coaching and tutoring were not controlled for. Furthermore, in this study, SAT and HSPA scores were the determinants of student achievement. The SAT measures verbal and mathematical reasoning skills, and the HSPA test focuses on Language Arts Literacy and Mathematics. While the HSPA test is tailored to assess whether eleventh grade students have the knowledge and skills required for graduation and is given to the majority of New Jersey public high school juniors, and the SAT test is the single most used standardized test of its type in the country the HSPA and SAT scores are not the only predictor of student achievement.

Additionally, this study was a causal-comparative study. A random assignment of groups was not utilized because the independent variable was already received. The study has weaknesses in its lack of randomization, manipulation, and control. A single empirical study of this kind cannot provide every answer, just some potential answers to the complex overarching question of what is the best pedagogy for teaching contemporary students.

This study utilized data from New Jersey School Report Cards and therefore, cannot hope to capture all of the characteristics of a school environment that might influence student learning. No other variables were studied. Academic achievement can be affected by curriculum, teacher age and education, and professional development. After reviewing the outcomes of the study the following recommendations are suggested for future research:

1. Replicate this study statewide at the middle school level. New Jersey has many middle schools involved currently in service learning. Although past studies have showe
that middle schools students are less impacted by service-learning, the dearth of new programs and greater awareness of the essential elements of service-learning may yield new results due to better implemented programs.

2. Replicate this study nationally by comparing several states. The National Youth Leadership Council has created a database of statistics from thousands of principals of middle and high school nationally regarding the school service-learning programs. A study that included states that rely heavily on SAT or other standardized test scores for achievement could be utilized for comparisons.

3. Create a qualitative study of service-learning impact that is longitudinal in design. For school year 2004-2005, eight new school districts in New Jersey received funding to implement service-learning programs school wide. A qualitative study could include triangulation from standardized scores past and present as well as personal testimonies.

4. Continue revisiting the Community Service-Learning and NJ Leader School districts and compare Grade Point Average, Advanced Placement scores or other performance indicators could do research regarding the current study.

5. A study focusing on the lower socioeconomic districts enabling researcher to further test and define what causes the greater academic impact.

6. Continued research into achievement by creating studies that utilize alternative measurements of achievement including portfolios, formative and summative evaluation by teachers, and rubrics.

By relating activities to real-life experiences, service-learning improves workplace readiness skills and personal development among students. Service-learning
gives students an opportunity to participate in the active-learning called for by Dewey nearly a century ago. As self-directed learners, students will continue to gain the important skills of collaboration, complex thinking and problem solving. Service-learning can rejuvenate old lessons and serve as a catalyst for new programs while encouraging life-long civic participation.
References


College Board, (2003). *College Board Tests.* Retrieved June 24, 04 from: 


Corporation for National Service, National Community Service Act of 1990.


*District Factor Grouping* (p.d.). Retrieved on August 02, 03 from http://www.state.nj.us/ njded/finance/st/dfgdesc.doc


at the annual Meeting of the American Educational Research Association.

Canada: Montreal.


What is service-learning? (n.d.) Retrieved from New Jersey Department of Education Website on September 02, 03 from: http://www.state.nj.us/njded/blsa/whatis.htm.


Appendix A

New Jersey Department of Education School Service Survey
Welcome to the NJ School Service Survey. This survey is designed to obtain vital information of service activities within your school. This brief survey will provide information that will be used to formulate future program and policy directions and will help us to be more responsive to schools needing assistance in the development of service programs. Please mail responses by Tuesday, November 26, 2002, to: The NJ Learn and Serve America Program, 100 River View Plaza, P.O. Box 500, Trenton, NJ 08625-0500. This form is available online at http://www.state.nj.us/njced/lsa If you have any questions, you may call Rowena Madden or Linda V. Rivers at 609-633-9627.

1. Does your school have a community service program(s)? ___ Yes ___ No
   If so, please describe the program(s) currently active in your school.

2. When does the community service occur? Please check all that apply:
   _____ During School Hours _____ After School Hours
   _____ During Summer Session _____ Not at All

3. What percentage of your student body is involved in service projects?
   _____ Less than 5% of the student body participates
4. On which areas do your service programs focus?

- ______ Beautification Projects
- ______ Community History
- ______ The Disabled
- ______ Environment
- ______ The Homeless
- ______ International
- ______ Literacy/Tutoring
- ______ Intergenerational
- ______ Young Children
- ______ Other (Please describe)

5. Do students receive academic credit for service activities?

- ______ Yes
- ______ No

6. Is service a graduation requirement, and if so how many hours per year are required?

- ______ Yes (Please check the appropriate number of hours below.)
- ______ 10 hours or less
- ______ 11-20 hours
- ______ 21-40 hours
- ______ 41-60 hours
- ______ 60 hours or above
- ______ No

7. Is there an evaluation given to your students for their service work?
8. Do students receive recognition for their service accomplishments and if so, what type?
   ______ Yes – Type: _______
   ______ No

9. Have service activities been integrated into the curriculum? If so, which classes?

10. How familiar are you with service-learning? Please check all that apply:
    ______ Do not knew what service-learning is
    ______ Heard of Service-learning but do not know very much about it
    ______ Would like more information on service-learning
    ______ Am very familiar with service-learning
    ______ Service-learning is no different from community service

11. Do you support the concept of service-learning based on the following definition?

   Service-learning is a method of instruction that combines meaningful service to the community with classroom studies in a way that improves learning and strengthens the community.
   ______ Yes
12. Does a service-learning program currently exist at your school?

_____ Yes

_____ No

13. What portion of your faculty/staff/administration is familiar with service-learning?

_____ Fewer than 25%

_____ More than 75% are familiar

_____ 25-50% are familiar

_____ Not Applicable

14. Is your district planning on implementing a service-learning program in the future?

_____ Yes

_____ No

_____ Am not sure

15. Would you support a state mandate (such as that of Maryland) making either community service or service-learning a graduation requirement for New Jersey Students?

_____ Yes

_____ No

16. Do you currently have a coordinator or director of service-learning or community service in your school? Who is responsible for managing your school's service work?
17. Who provides transportation for off-campus student service activities. Please select as many as apply.

- School provides at least some transportation
- Parents provide transportation
- Students provide own transportation
- Transportation provided by all of the above: school, parents and students

18. How has your school addressed the issue of liability for off-campus student service activities?

19. How are your service activities funded? Please select as many as apply.

- School Boards provide funding for salaries & transportation
20. Do you use service activities to provide students with opportunities to explore career options?

___ Yes
___ No

22. What impact has involvement in community service/service-learning had on your students? Please select as many as apply.

___ Promotes good citizenship and character development
___ Helps solve community problems
___ Develops leadership and teamwork skills
___ Gives students an opportunity to feel they have made a difference
___ Serves as a good vehicle for student career exploration
___ Improves students' academic performance
___ Increases the likelihood of volunteering as adults
__ Provides good community relations for the school
__ Fosters inter-generational relationships within the community
__ Little, if any, impact
__ Difficult to administer
__ Do not have enough information

Thank you for completing this survey!

Please send completed copies to:
Leadership and Vision Program, 100 Milrond Plaza, P.O. Box 500